

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 09:19:25 ; Search time 1703 Seconds

(Without alignments)
8207.108 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863
Sequence: 1 gcacgagctctacacacac.....aaaaaaaaaaaaaaaaaaaa 863

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 60

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	65.1	562	14	BM732121 sal73b11.
2	370	42.9	586	12	BB820766 GM700012A
3	349	40.4	431	10	AA459279 sn22h08.y
4	285	33.0	450	10	BE058947 sn23b01.y

ALIGNMENTS

RESULT 1
LOCUS BM732121
DEFINITION sal73b11.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl061-4678 5' similar to TR:Q39028 Q39028 ATMYB2.1; mRNA sequence.
ACCESSION BM732121
VERSION BM732121
KEYWORDS BM732121.1 GI:19053454
SOURCE EST.
ORGANISM soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 562)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schuk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Glibco
High quality sequence stop: 421.

FEATURES

source

Location/Qualifiers

1..562
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl061-4678"
/clone_lib="Gm-cl061"
/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/note="Vector: pluscript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluscript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT

187 a 115 c 128 g 132 t

Query Match 65.1%; Score 562; DB 14; Length 562;

Best Local Similarity 100.0%; Pred. No. 3.3e-128; Mismatches 0; Indels 0; Gaps 0;

Query	Match	Score	DB	Length	562;
OY	10	CTATCACACACAAAGTCAATGATATAAAACAAGTGTAAAGCTCTCAAGATCTGA	69		
DB	1	CTATCACACACAAAGTCAATGATATAAAACAAGTGTAAAGCTCTCAAGATCTGA	60		
OY	70	AGTGAGAAAGGCGCTTGAGCAATGAGCAAGACCTTGATCTTGATGAACATATATGCAA	129		

ACCESSION	FACTOR ; mRNA sequence.
AM459279	
KEYWORDS	AM459279.1 GI:7029496
SOURCE	EST.
ORGANISM	soybean.
REFERENCE	Glycine max
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine. 1 (bases 1 to 431)
TITLE	1 (bases 1 to 431)
JOURNAL	Shoemaker,R., Keim,P., Vodka,L., Erpellding,J., Corwell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,B., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
COMMENT	Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: east@watson.wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: cculeresgen.com Insert Length: 723 Std Error: 0.00 Seq Primer: -40RP from Glbco High quality sequence stop: 411. Location/Qualifiers 1..431 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-5560" /clone_id="Gm-cl016" /tissue_type="immature flowers of field grown plants" /lab_host="XLI0-Gold" /note="Vector: plasmidscript II XR; Site1: EcoRI; site2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene plasmidscript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the plasmidscript vector. The ligated cDNA fragments were transformed into XLI0-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpellding."
BASE COUNT	145 a 81 c 88 g 117 t
ORIGIN	
Query Match	40.4%; Score 349; DB 10; Length 431;
Best Local Similarity	100.0%; Pred. No. 3.3e-76;
Matches 349; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	336 CCTGAAGACTGATTAATGATCAAGACTTTTGAGAGACAAGATCCAGAAGCATTC 385
DB	 8 CCGTGAAGACGTAAATGATCAAGAACTATTGGAGAGACAAGATCCAGAAGCATAAC 67
OY	386 AAGCAAGCTGAGAACTTTCAGACAGAGTAGTAAATTAATTCGAGATAATGATCACCAA 445
DB	 68 AAGCAAGCTGAGAACTTTCAGACAGAGTAGTAAATTAATTCGAGATAATGATCACCAA 127
OY	446 GCTAGCACTAGGCATGTTTCACACATGGCTGAGGCCCATGAGAGATGATTCTCCACCCTGT 505
DB	 128 GCTAGCACTAGGCATGTTTCACACATGGCTGAGGCCCATGAGAGATGATTCTCCACCCTGT 187

QY	506	TATCAAGGAATGTTGAGGCAATTTTCACACTAGTTCCTCAATATATCTGATCAATCC	565
Db	188	TATCAAGGAATGTTGAGGCAATTTTCACACTAGTTCCTCAATATATCTGATCAATCC	247
QY	566	AGTTGTTGACCAATGACACAAACAACCTTAACTATTTGGAGCATGAGCATAGCTGTCTCA	625
Db	248	AGTTGTTGACCAATGACACAAACAACCTTAACTATTTGGAGCATGAGCATAGCTGTCTCA	307
QY	626	ATGCATTTACTGACGCTGATTAATATATATCAAGATAAACCAAGTT	674
Db	308	ATGCATTTACTGACGCTGATTAATATATATCAAGATAAACCAAGTT	356
RESULT 4			
LOCUS	BE058947	450 bp	MRNA
DEFINITION	sn2b01.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl016-12290 5' similar to TR-092TDS 092TDS PUTATIVE TRANSCRIPTION FACTOR ; mRNA sequence.		
ACCESSION	BE058947		
VERSION	BE058947.1	GI:8403313	
KEYWORDS	EST.		
SOURCE	soybean.		
ORGANISM	Glycine max		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
AUTHORS	1 (bases 1 to 450) Shoemaker,R., Kelm,P., Vodkin,L., Expelling,J., Corvett,J., Khanna,A., Bolla,B., Maric,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Rittler,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R. Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R./Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Hunttsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccr@resgen.com Seq primer: -40RP from Gibco High quality sequence stop: 411.		
TITLE	Public Soybean EST Project		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Shoemaker R./Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Hunttsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccr@resgen.com Seq primer: -40RP from Gibco High quality sequence stop: 411.		
FEATURES	Location/Qualifiers		
SOURCE	1..450 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl016-12290" /clone.lib="Gm-cl016" /tissue.type="Immature flowers of field grown plants" /lab.host="Xl10-Gold" /note="Vector: pluescript II XR, Site.1: EcoRI, Site.2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapter were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into Xl10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Expelling."		
BASE COUNT	151 a	84 c	90 g 125 t
ORIGIN			

Query Match 33.0%; Score 285; DB 10; Length 450;
Best Local Similarity 99.7%; Pred. No. 1.2e-60;
Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	339	ATAATGAGATCAGAGAACTATTGGAGACAGATCCAGAAAGACATCAAGCAGCTGAGA	398
Db	36	ATAATGAGATCAGAGAACTATTGGAGACAGATCCAGAAAGACATCAAGCAGCTGAGA	95
QY	399	ACTTTCAGCAGAGAGTAGTATATCTGAGATAANTGATCAGCAAGCTAGCACTAGCC	458
Db	96	ACTTTCAGCAGCAGAGTAGTATATCTGAGATAANTGATCAGCAAGCTAGCACTAGCC	155
QY	459	ATGTTCCACCATGGCTGAGCCCATGGAGATGTAATCTCCACCCGTATATCAAGGAATGT	518
Db	156	ATGTTCCACCATGGCTGAGCCCATGGAGATGTAATCTCCACCCGTATATCAAGGAATGT	215
QY	519	TAGAGCCATTTTCAACTAGTTCCTCAATTAATTCCTGATCAATCCAGTTGTGTACCA	578
Db	216	TAGAGCCATTTTCAACTAGTTCCTCAATTAATTCCTGATCAATCCAGTTGTGTACCA	275
QY	579	ATGACACACACACATTAATTAATTTGGAGCATGGAGATAGCTGTCATGCAATTTACTGA	638
Db	276	ATGACACACACACATTAATTAATTTGGAGCATGGAGATAGCTGTCATGCAATTTACTGA	335
QY	639	ACGGTGATTAATATATATCAAGATAAAACCTAAGTT	674
Db	336	ACGGTGATTAATATATATCAAGATAAAACCTAAGTT	371

Search completed: February 18, 2003, 10:31:23
Job time 1705 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:08:59 ; Search time 2178 Seconds

(without alignments)
11531.553 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 1 gcacgagcctctacacac.....aaaaaaaaaaaaaaaaaaaa 863

Sequence: 1 gcacgagcctctacacac.....aaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382.8	44.4	1022	8	PSMYB26
2	239.4	27.7	910	8	AY086615 Arabidops
3	236.2	27.4	645	8	AF175987 Arabidops
4	215	24.9	1033	8	AB005888 Arabidops
5	208	24.1	1127	8	AY088761 Arabidops
6	194.6	22.5	872	8	AB058642 Arabidops
7	176.6	20.5	1117	8	AF510112 Craterost
8	175	20.3	1295	8	CPU33917 Arabidops
9	163.4	18.9	1423	8	AF262733 Arabidops
10	156.4	18.1	1343	8	AY026332 Arabidops
11	155.6	18.0	1140	8	AF334815 Arabidops
12	154	17.8	817	8	AF133705 Arabidops
13	149.2	17.3	931	8	AY008377 Arabidops
14	149.2	17.3	1024	8	AF062894 Arabidops
15	146.2	16.9	1033	8	AF034133 Gossypium
16	144	16.7	133242	8	AF003140 Oryza sat
17	144	16.7	145491	8	AP002883 Arabidops
18	142.2	16.5	771	8	AF272733 Arabidops
19	138	16.0	1753	8	AF411970 Arabidops
20	137.8	16.0	1070	8	AT062743 Arabidops
21	135.2	15.7	1337	8	LETHM16 Arabidops
22	134	15.5	1373	8	LETHM1 Arabidops
23	131	15.2	1341	8	AY099777 Arabidops
24	131	15.2	1364	8	AB005889 Arabidops
25	130.8	15.2	844	8	AT026397 Arabidops
26	130.4	15.1	1373	8	AF214116 Arabidops
27	130	15.1	1293	8	AY062619 Arabidops
28	129.6	15.0	984	8	AF175990 Arabidops
29	128.6	14.9	1038	8	AF249309 Arabidops
30	128.6	14.9	1493	8	AF062916 Arabidops
31	128.4	14.9	1335	8	ATCM4GENE Arabidops
32	128	14.8	1132	8	AY096523 Arabidops
33	128	14.8	1512	8	AY063939 Arabidops
34	127.6	14.8	1262	8	AF371980 Arabidops
35	127.2	14.7	1220	8	AF161711 Arabidops
36	127	14.7	981	8	AF370613 Arabidops
37	127	14.7	1252	8	ATM1X1A Arabidops
38	126.8	14.7	1331	8	AY087252 Arabidops
39	126.6	14.7	1513	8	AF427146 Arabidops
40	126.6	14.7	1601	8	ZM057002 Arabidops
41	126.6	14.7	1802	8	MZEBPR Arabidops
42	126.2	14.6	1606	8	AF034132 Arabidops
43	125.2	14.5	318	8	AF474125 Arabidops
44	125.2	14.5	579	8	AF502295 Arabidops
45	124.4	14.4	813	8	AF175993 Arabidops

ALIGNMENTS

RESULT 1
PSMYB26
LOCUS PSMYB26 1022 bp mRNA linear PLN 09-FEB-1998
DEFINITION P sativum mRNA for Myb-like protein (Myb26).
ACCESSION Y11105
VERSION Y11105.1 GI:1841474
KEYWORDS myb-like protein; Myb26.
SOURCE Pium sativum
ORGANISM Pium sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae;
Pium.
REFERENCE 1 (bases 1 to 1022)
Umarl, A. and Strommer, J.

Db	668	TGTCGAATGACACAAATTAATTTGGAGCATGAGATATCTGTCAATGCA	727
QY	632	TTACTGAGCGATTAAT 651	
Db	728	TTACTCAATGAGCATTTAAT 747	
Result 2			
LOCUS	AY086615	910 bp	mrna linear PLN 25-JUN-20020
DEFINITION	Arabidopsis thaliana clone 262460	mrna	complete sequence.
ACCESSION	AY086615		
VERSION	AY086615.1	GI:21405325	
KEYWORDS	FLI CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Esakaya, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustroids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 910)		
REFERENCE	Haas,B.J., Volkofsky,N., Town,C.D., Troukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.		
AUTHORS	Full-length messenger RNA sequences greatly improve genome annotation		
TITLE	Genome Biol. (2002) In press		
JOURNAL	2 (bases 1 to 910)		
REFERENCE	Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
AUTHORS	Full-length cDNA from Arabidopsis thaliana unpublished 3 (bases 1 to 910)		
TITLE	Broyer,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA		
AUTHORS	This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Luer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genes carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5 sequences, selection of clones, and sequence assembly.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..910		
	/organism="Arabidopsis thaliana"		
	/db_xref="taxon:3702"		
	/clone="262460"		
	110..754		
CDS	/codon_start=1		
	/product="putative transcription factor MYB24"		
	/protein_id="AA063674.1"		
	/db_xref="GI:21554734"		
	/translation="MEKRSGGSGSGDAEVRKGPMTMEDLLINIANHGEYWNLSAKSGLRTRGSRRLNLTLPDYRGRNITPEEDLTMEIHLAKWNRMSITAEHLGLTDEINFWRTIKRYLIKSGEYTWVGSSSEFINHAATSHVANDIOETMDYSSITSYOHAANINQINQYVDESSIMPLSDVSEONWYSDVIMPMINYNGN"		
BASE COUNT	307 a 169 c 185 g 249 t		
ORIGIN			
Query Match	27.7%	Score 239.4	DB 8: Length 910;
Best Local Similarity	72.4%	Pred. No. 1,1e-41;	
Matches 325; Conservative	0;	Mismatches 121;	Indels 3; Gaps 1;

[illegible]

LOCUS	AF175987	645 bp	mRNA	linear	PLN 30-AUG-2001
DEFINITION	Arabidopsis thaliana putative transcription factor (MYB24) mRNA,				
ACCESSION	complete cds.				
VERSION	AF175987				
KEYWORDS	AF175987.1	GI:5823306			
SOURCE	.				
ORGANISM	Arabidopsis thaliana.				
	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE	1 (bases 1 to 645)				
AUTHORS	Kranz,H.D., Denekamp,M., Greco,R., Jin,H., Leyva,A., Weissner,R.C.,...				
	Petroni,K., Urzainqui,A., Beyer,M., Martin,C., Smeekens,S., Tonelli,C., Paz-Ares,J. and Weisshaar,B.				
	Towards functional characterisation of the members of the R2R3-MYB gene family from Arabidopsis thaliana				
TITLE	Plant J. 16 (2), 263-276 (1998)				
JOURNAL	MEDLINE				
PUBMED	99056848				
	9839469				
REFERENCE	2 (bases 1 to 645)				
AUTHORS	Stracke,R., Werber,M. and Weisshaar,B.				
TITLE	The R2R3-MYB gene family in Arabidopsis thaliana				
JOURNAL	Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)				
MEDLINE	2181677				
PUBMED	11597504				
REFERENCE	3 (bases 1 to 645)				
AUTHORS	Stracke,R. and Weisshaar,B.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-AUG-1999) Dept. Biochemie, Max-Planck-Institut fuer				
	Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany				
FEATURES	Location/Qualifiers				
source	1..645				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
	/note="ecotype: Columbia"				
	1..645				
gene					

CDS	ORIGIN
<code> /gene="MYB24" 1..645 /gene="MYB24" /note="R2R3-MYB family member" /codon_start=1 /product="putative transcription factor" /protein_id="AAD53092.1" /db_xref="GI:5823307" /translation="MKRRSSGSGSGSDAIVKCPMTMEDLLILNTIANHGSGVYNS LAKSAGLAKRTGKSCRLRLMNLRLPDRVKRGNTTPEQLIMPLAKMGNRNSKIALHLP GRTDSLNKFNRIKIQRILIKSGETTVGSQSSFIINHATSTHVNMDTDETDWYSP TTTQVHSAINSINQOLNLYNGYVPESSIMPLSLVDSQSEONTWSYVDLMPNMYNGN" </code>	<p>BASE COUNT</p> <p>219 a 125 c 153 g 148 t</p>

Query Match	27.4%	Score 236.2	DB 8	Length 645
Best Local Similarity	71.9%	Prod. No. 5.8e-41		
Matches 323	Conservative	0	Mismatches 123	Indels 3
			Gaps	1
OY	56	TCTCATGATCCCTGAAGTGAAGAAAGGCGCTTGAGCAATGGAAGAAAGACTTGATCTGTATG	115	
Db	34	TCAGAGAGATGCGAGAGTGAAGAAAGGCGCATGAGCAATGGAAGAAAGATTTGATTTCTATC	93	
OY	116	AACTTATTGCAATCATGAGGAGGTGTTTGAACCTTTGGCCAAAGCTGCTGCTC	175	
Db	94	AATTATTATCGCCAAATCATGATGTGAAAGTGTGGAAGCTCTCGCCAAATCTGCAGAGACTA	153	
OY	176	AAACCTAACCGGAAAGATGGCCGGCTTAAGGAGGCTAAATTTACCTCCGCTCGATGTTTGA	235	
Db	154	AAAGCCACCGGAAAGAAAGTGGCCGGCTCGGGTGGCTGAACTTACCTCCGACTGATGTGCGA	213	
OY	236	AGAGGGAATATTACACCCGAGAACACACTTTTGAATTATGAGAGCTCCAGCAAAAGTGGGA	295	
Db	214	CGGGAAATATTCACACCCAGGAAGAACGCTTCACCATCATGAGAACTTCATGAAATATGCGGA	273	
OY	296	AACAGGTGTCCTCAAAATTTGGCCAGCATCTACCTGGAAGAGATGAATAGATGCAAGAC	355	
Db	274	AATAGTGTGTCAAAATTTGCAAAAGATTTTACCAGAGAAAGCCAGCATGAGATTAAAGAT	333	
OY	356	TATTGAGAGACAAAGATCCAGAACGACATCAAGCAAGCTGAGAACTTTTCAGCAACAGAGT	415	
Db	334	TTTTGGAGAGACTAAGATCCAGAAATATCATCATCATCAAGACGGGAAACGACGCGTTGGA	393	
OY	416	AGTATTATTTCGAGATTAAGATGATACCAAGATAGCATAGCACTGTTCCACATGGCT	475	
Db	394	TCACAAAGCTCCGAGTTTATTAACCATATGATGAGCAACGACCATGATGATGATGATGAT	453	
OY	476	--GAGCCCATGAGAGATGATTTCTCCACC	501	
Db	454	CAGAAACCATGAGATGATGATTTCTCCAAC	482	
RESULT 4				
AB005888				
LOCUS	AB005888	1033 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana mRNA for ATMYB3, complete cds.			PLN 05-FEB-1999
ACCESSION	AB005888			
VERSION	AB005888.1	GI:2280527		
KEYWORDS	ATMYB3			
SOURCE	Arabidopsis thaliana (strain Columbia)		cdna	to mRNA.
ORGANISM	Arabidopsis thaliana			
REFERENCE	Noti,M., Urao,T., Shinzaki,K.Y. and Shinzaki,K.			
AUTHORS	Molecular cloning of three cDNAs encoding novel myb homologs from			
TITLE	Arabidopsis (Accession Nos. AB005888 and AB005889) (FGR98-111)			
JOURNAL	Plant Physiol. 117, 720 (1998)			
REFERENCE	Noti,M.			
AUTHORS	Direct Submission			
TITLE	Submitted (18-JUL-1997) Masaki Noti, Chiba University, Faculty of			

LOCUS AB058642 872 bp mRNA linear PIN 04-APR-2001
DEFINITION Lilium hybrid division I LmYb mRNA, complete cds.
ACCESSION AB058642
VERSION AB058642.1 GI:13537529
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Nakatsuka, A., Izumi, Y. and Yamagishi, M.
Isolation and characterization of the genes related to anthocyanin biosynthesis in Asiatic hybrid lily
Unpublished
2 (bases 1 to 872)
Nakatsuka, A., Izumi, Y. and Yamagishi, M.
Direct Submission
Submitted (24-MAR-2001) Akira Nakatsuka, Shimane University, Faculty of Life and Environmental Sciences, Nishikatsu 1060, Matsue, Shimane 690-8504, Japan
(E-mail:nakita@life.shimane-u.ac.jp, Tel:81-852-32-6502)
Location/Qualifiers
1..872
/organism="Lilium hybrid division I"
/cultivar="Montreux"
/db_xref="taxon:156532"
/tissue_type="tepals"
1..872
/gene="LmYb"
80..634
/gene="LmYb"
/codon_start=1
/protein_id="BAB40790.1"
/db_xref="GI:13537530"
/translation="MDKRGVTEDEVKRRGPTWMEDELILINIANHGDGVNMTLARS
AGLKRKSGCSRLRMVLPDYRRNGITPEBOLLIMELHDPWGNWRSKIAELPERD
NEIKMPRTYRQKSKKNSESSFGQISMLSDETSAQMEDVDVGSQTSYDGGQTSNNL
DAFEIPEPBDMLSVIEDMAQY"
BASE COUNT 269 a 177 c 213 g 213 t
ORIGIN
Query Match 22.5%; Score 194.6; DB 8; Length 872;
Best Local Similarity 69.1%; Pred. No. 5,6e-32;
Matches 266; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 52 GACGTCACAGATCTGAAGTGAAGAAAGGCTTGACAAATGGAAGAGACTTGATCTT 111
111 11111 111 11111 111 11111 111 11111 11 11111
DB 97 GACTATGGAAGATGTGAGGTGAGAAAGACATGACCATGGAAGAGATCTCATCT 156
QY 112 GATGAACATATATGCAAAATCATGCGGAGAGGTGTTGAACTCTTTGGCCAAAGCTGCTG 171
111 11111 11 11111 11 11111 11 11111 1111 1111
DB 157 CATCAACTACATAGCCACACATGCGACGAGCTGTGAACACAGCTGCTGATCGGCCG 216
QY 172 TCTCAACGTAACGGAAGAGTGGCGCTTAAGGTGGCTAAATTAATTCCTCCGCTGATAT 231
111 11 11111111 111 1111 1111111111 11111111 11111111
DB 217 ACTGAAGAGAGACTGGAAAGAGCTGAGGCTGTGCTGTAATTAATTCAGAGCTGATAT 276
QY 232 TAGAAGAGGAATATTAACACCCGAGCAACACTTTTGAATTAAGAGCTTCACGCAAAAGT 291
111 11111 11111 11111 1111 1111 11111111 1111 11111
DB 277 TCGTCGGGTAAACATCACACAGAGAGACAGCTTGTGATCATGGAACCTTCACAGCAGAG 336
QY 292 GGGAAACAGGTGTCGCAAAATGTCACAGACTTACCTGGAAGAGCTGATTAATGAGATCA 351
111 11111 11111 11111 111 11 1111 11 1111 11111111
DB 337 GGGTAACAGATGTGCAAAATTTGCAAAAGAGTCCACAGACCAACAGTAAGCAATCA 396
QY 352 GAACATTTGAGAGCAAGAGTCCAGAACATCAAGCAAGCTGAGAACTTCAGCAACA 411
1111 1111111 11 1111111 111111 11 1111 11 111111
DB 397 GAACCTTCTGAGAGACAGAGTCCAGAAATTCAGAAATTCGCGAATCACTGAGGCCA 456
QY 412 GAGTAAATTAATTTGAGATTAAT 436
111 11 1111 1111
DB 457 AATCTCCATGCTTTGGATGAACCT 481
1111 1111 1111

RESULT 7
LOCUS AF510112 1117 bp mRNA linear PIN 26-MAY-2002
DEFINITION Craterostigma plantagineum MYB transcription factor (MYB10) mRNA, complete cds.
ACCESSION AF510112
VERSION AF510112.1 GI:21215201
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
Nakatsuka, M.A., Bartels, D. and Iturriaga, G.
The Craterostigma plantagineum CPMYB10 gene enhances stress tolerance to Arabidopsis transgenic plants
Unpublished
2 (bases 1 to 1117)
Villalobos, M.A. and Iturriaga, G.
Direct Submission
Submitted (07-MAY-2002) Plant Molecular Biology, Institute of Biotechnology, Universidad Ave. 2001, Cuernavaca, Mor 62210, Mexico
Location/Qualifiers
1..1117
/organism="Craterostigma plantagineum"
/db_xref="taxon:4153"
1..1117
/gene="MYB10"
/note="CPMYB10"
79..1080
/gene="MYB10"
/codon_start=1
/product="MYB transcription factor"
/protein_id="YAM43912.1"
/db_xref="GI:21215202"
/translation="MNOQOYVSKRNKRVNNCEDDDSDLRGRGPTWVDEPTLLIN
YAHGEGRMNSLAREFAGLRKRGKSCRLRLVLPDYRRNGITPEBOLLIMELHDSRGN
RMSKIAHLBGRITNEIKNTYRTPVQVHAQOLKDVNSKPKFTMRILMPRIYERLQ
ASATTTDGAAPPAVASPSSAMNTRACYSAAQAAGDHRQFMLMPOYATATTTNNSMI
AQENSTVASSEFSGSLSELTAAANYANTRYINGADHOCIDSTSYDMQCAVGN
NGNDOLGIMGADDRSRNEDQMMWMTTDDVYNGSSDQDNLMMVVDVWFIQPSSCP"
BASE COUNT 329 a 265 c 271 g 252 t
ORIGIN
Query Match 20.5%; Score 176.6; DB 8; Length 1117;
Best Local Similarity 71.3%; Pred. No. 4,2e-28;
Matches 233; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 64 TCCCTAAGTGAAGAAAGGCTTGACAAATGGAAGAGACTTGATGATGACTATAT 123
111 11 11111 111 1111111 11 1111111 11 1111111
DB 150 TTGGAACCTCAGAAAGAGACTTGAGACTGTGATGAACATCACTCACTCACTCAT 209
QY 124 TGCAAATCATGCGGAAGGTGTTTGAACCTTTTGCCAAAGCTGCTGTCAAAGTAA 183
111 11111 11111 1111111 11 11111 11111 1111111
DB 210 CGCTCACCATGCGGAAGGAAGATGGAACCTCTTGACAGCTTTCCTGCGCCGAAACGAC 269
QY 184 CGGAAGAGTTGCCGCTTAAGTGTCTAAATTAATTAATTAATTAATTAATTAATTA 243
1111111 111 11 11111 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 270 TCGAAAGAGCTGCACATTTGATGATGAGTCCAGCAAGTGGGGAACAGAGT 329
QY 244 TATTTACACCCGAGCAACATTTTGAATTAATGAGTCCAGCAAGTGGGGAACAGAGT 303
111 1111 1111 1111 1111 111 11 1111 11 11111 11 111111
DB 330 TATTCACATCGAAGAACAGCTTTCATTTCTTAATTCATTCACCTTGAGGCAATAGGTG 389
QY 304 GTCCAAAATTTGCCAAGACTTACCTGGAAGAGCTGATATGATGATCAAGAACTTTGGAG 363
111 11 11111 11 111111111 11 1111111 111111 111111
DB 390 GTCCGAAGATTTGCCCAACATTTTGCCTGGAAGAGACGACAAATGATTAAGAATCTGTGAG 449
111 11111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

Db 450 AACAGAGGTCCAAAACATGCCAACA 476

RESULT 8

CPU33917 1295 bp mRNA linear PLN 17-Apr-2001

LOCUS Craterostigma plantagineum myb-related transcription factor (cpm7)

DEFINITION mRNA, complete cds.

ACCESSION U33917

VERSION U33917.1 GI:1002799

KEYWORDS Craterostigma plantagineum.

SOURCE Craterostigma plantagineum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Lamiales; Lamiales incertae sedis;

Toreniaeae; Craterostigma.

1 (bases 1 to 1295)

REFERENCE Iturriaga, G., Leyna, L., Villegas, A., Charalbeh, R., Salami, F. and

Barthelemy, D.

A family of novel myb-related genes from the resurrection plant

Craterostigma plantagineum are specifically expressed in callus and

roots in response to ABA or desiccation

Plant Mol. Biol. 32 (4), 707-716 (1996)

JOURNAL MEDLINE

97134962

8980522

2 (bases 1 to 1295)

REFERENCE Iturriaga, G.

Direct Submission

Submitted (14-AUG-1995) Gabriel Iturriaga, Plant Molecular Biology,

Instituto de Biotecnologia, Av. Universidad #2001, Cuernavaca, MOR

62210, Mexico

FEATURES

Location/Qualifiers

1..1295

/organism="Craterostigma plantagineum"

/db_xref="taxon:4153"

gene

1..1295

/gene="cpm7"

125..1132

/gene="cpm7"

/note="putative DNA-binding protein; Description: myb-like

gene; myb-related transcription factor"

/product="Cpm7"

/codon_start=1

/protein_id="AAB58314.1"

/db_xref="GI:1002800"

/translation="MNQOYVYKSKNNKQINSEDDSDLRGPGWTVDEPFLINITYA

HGEGRWNSIARFAGLKRKTSQRLMINTLRDVRGRNTLLEQILLILSHRMGRN

WSKIAQHLPGRTNENIKNTYRTYQKHAQKCDVNSKQPKDTRILMPRIYERIOR

AAITDDGAPPLASSASPSAMNTACTSTMAAGDRHROPMLPQYIAMTTHSNLITA

QENSTYVASESEFGLSSSELTEANVANYHOVINGADHQIDSSSTTSYGVQNDPAINT

VGNNGNSDIQMGMSVDDRSNEDQMMMDVDNGSSDKDNLIMVNDVWFLDOOFSSC

F"

BASE COUNT 405 a 286 c 279 g 325 t

ORIGIN

Query Match 20.3%; Score 175; DB 8; Length 1295;

Best Local Similarity 70.9%; Pred. No. 9.1e-28;

Matches 232; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 64 TCCTGAGTGAAGAAAGGGCTTGACAAATGGAAGAACTTGATGAACTATAT 123

DB 193 TTGGAAGCTCGAAGAGAGACCGTGACTGTGATGAAGACTTCATCACTACAT 252

QY 124 TGCAAAATATGGGAAGAGTGTGGAACCTCTTGCCAAAGCTCTGCTCAAAAGCTAA 183

DB 253 CGCTCACCATGCGAAGAGAGATGAACTCTCTGACAGTTTCCCTGCGCTGAAGACGAA 312

QY 184 ZGGAAGAGCTTCCGCGCTAAGTGGCTAAATTAACCTCGCTGCTGATGAAAGAGGAA 243

DB 313 TGAAGAAGCTGCAAGATGATGATTAACACTACTTGAGACCCGATGTGCTGAGGAAA 372

QY 244 TATTACACCCGAGAGAACACTTTGATTATGAGCTCCAGCAAAAGTGGGGAACAGGTG 303

DB 373 TATCACACTCGAAGAACACACTCTCTCATTCATTCATTCCTGCGGCAATAGCTG 432

QY 304 GTCCAAATTTGCCAGACACTACCTGAGAGACTGATATGATGATCAATCAATATTTGGG 363

DB 433 GTGGAAGATTGCCACACATTTCCTGGAAGAGACAGACAAATGATTAAGAACTACTGGAG 492

QY 364 GACAGGATCCAGACAGCATCAAGCA 390

DB 493 AACGAGGTCCAAAACATGCCAACA 519

RESULT 9

AF262733 1423 bp mRNA linear PLN 30-AUG-2001

LOCUS Arabidopsis thaliana putative transcription factor MYB108 (MYB108)

DEFINITION mRNA, complete cds.

ACCESSION AF262733

VERSION AF262733.2 GI:15375290

KEYWORDS Arabidopsis thaliana.

SOURCE Arabidopsis thaliana

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1423)

REFERENCE Stracke, R., Werber, M. and Weisshaar, B.

The R2R3-MYB gene family in Arabidopsis thaliana

Curr. Opin. Plant Biol. 4 (5), 447-456 (2001)

JOURNAL MEDLINE

21481677

11597504

2 (bases 1 to 1423)

REFERENCE Stracke, R. and Weisshaar, B.

Direct Submission

Submitted (02-MAY-2000) Blochemie, Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany

3 (bases 1 to 1423)

REFERENCE Stracke, R. and Weisshaar, B.

Direct Submission

Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,

Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg

10, Koeln D-50829, Germany

Sequence update by submitter

On Aug 30, 2001 this sequence version replaced gi:8101955.

FEATURES

Location/Qualifiers

1..1423

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="3"

/note="ecotype: Columbia"

1..1423

/gene="MYB108"

1..154

/gene="MYB108"

5

/gene="MYB108"

/note="alternative transcription start site"

155..1126

/gene="MYB108"

/note="At3g06490; member of the R2R3-MYB family"

/product="putative transcription factor MYB108"

/codon_start=1

/protein_id="AA072668.1"

/db_xref="GI:8101956"

/translation="WDEGRSLKNNMEDNDLKRGPWTAEDFKLNIYATNGEGRN

NSLSKAGLQRTGKSCRLRWLNLKPDVRRGNTLLEQLILLELSRWGNWSKIAQ

LEGRDNEIKNTYRTYQKHAQKCDVNSKQPKDTRILMPRIYERIOR

AAITDDGAPPLASSASPSAMNTACTSTMAAGDRHROPMLPQYIAMTTHSNLITA

QENSTYVASESEFGLSSSELTEANVANYHOVINGADHQIDSSSTTSYGVQNDPAINT

VSADLTETYSAPNPNPAYSGOMKNSYTPPOINTVSSQLLEPDNFDYRGLDDEDTAM

QOSNLSWNEENI NGAASSSDSLMNIWIGETDEEFWFLQDQDQGNNGSG"

1127..1423

/gene="MYB108"

3' UTR

misc_feature

5' UTR

CDs

[illegible]


```

VERSION      AY008377.2  GI:15375307
KEYWORDS
SOURCE       Arabidopsis thaliana.
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE    1 (bases 1 to 931)
AUTHORS      Stracke,R., Werber,M. and Weisshaar,B.
JOURNAL      The R2R3-MYB gene family in Arabidopsis thaliana
MEDLINE      21481677
PUBMED       11597504
REFERENCE    2 (bases 1 to 931)
AUTHORS      Stracke,R. and Weisshaar,B.
JOURNAL      Direct Submission
TITLE        Submitted (29-SEP-2000) Dept. Biochemie, Max-Planck-Institut fuer
              Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln D-50829, Germany
REFERENCE    3 (bases 1 to 931)
AUTHORS      Stracke,R. and Weisshaar,B.
JOURNAL      Submitted (29-AUG-2001) Dept. Plant Breeding and Yield Physiology,
              Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg
              10, Koeln D-50829, Germany
REMARK       Sequence update by submitter
COMMENT      On Aug 30, 2001 this sequence version replaced gi:11641119.
FEATURES
             source
             1..931
                /organism="Arabidopsis thaliana"
                /db_xref="taxon:3702"
                /chromosome="1"
                /note="ecotype: Columbia"
             1..83
             6
                /note="alternative transcription start site"
             84..815
                /note="Atlg48000; R2R3-MYB family member"
             /codon_start=1
             /product="putative transcription factor MYB112"
             /protein_id="AG38379.2"
             /db_xref="GI:15375308"
             /translation="MNISRTFANCKTLINHKKEVEVEKKEIEIRSGPMTVEEDMK
              LVSYSILHGEGRWNSLSRSAGLNRTGSCRMLANLRPDIRGDSKOKPDRIKIMMPRL
              SRMGNSRSKIACHLPGRTPDEIKYKWRTPQKAKLLKCDYNSKOKPDRIKIMMPRL
              IERTAAIOSVOTFSNMHSPEPNSVATATSTSSSEAVRSFYGDDVEGTLDHMTNG
              GYWPNGGDTEFTLCSFDELNKLWLIQ"
             816..931
             897
                /note="alternative"
             897
                /note="polyA_site"
BASE COUNT   317 a      183 c      194 g      237 t
ORIGIN
Query Match      17.3%; Score 149.2; DB 8; Length 931;
Best Local Similarity 67.2%; Pred. No. 3.7e-22;
Matches 211; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

```

DB 414 AAGATTGCTCAACATTTACCGGGAGAGACAGATACGATTAAGATTATTGGAGACA 473
OY 368 AGATCCGAAGCA 381
DB 474 CGCGTTCAAAAGCA 487
FEATURES
             source
             1..1024
                /organism="Arabidopsis thaliana"
                /cultivar="Columbia"
                /db_xref="taxon:3702"
             1..1024
             /gene="MYB59"
             /allele="COL"
             70..774
             /gene="MYB59"
             /note="MYB59; R2R3-MYB factor family member"
             /codon_start=1
             /product="putative transcription factor"
             /protein_id="FAC83616.1"
             /db_xref="GI:3941480"
             /translation="MKLYOERYKGRPTQEDLLVNFVHLGDRMDVVAVSGINR
              TGKSCRLLWVNYLHPGLKRGKMTTPQERLVLELHAHWKNSKILAKLGRDNEIKN
              YWRTHMRKKADEKRRPMSPTSSSNCCSSMTTQDYGSGNKGKNOEGEGYSMD
              DIMREIDOSGHVIRPVADNYNSREOSCLNPPPLASPTWSSLESISINWMDAESKMSF
              AIDQPLSPFEGHSGL"
BASE COUNT   344 a      183 c      230 g      267 t
ORIGIN
Query Match      17.3%; Score 149.2; DB 8; Length 1024;
Best Local Similarity 67.2%; Pred. No. 3.6e-22;
Matches 211; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

QY 248 ACACCGAGAACAACTTTGATTATGAGCTCCAGCAAGTGGGAAACAGGTGCTCC 307
AF034133 1033 bp mRNA linear PLN 11-JUN-2002
LOCUS Gossypium hirsutum MYB-like DNA-binding domain protein (MYB5) mRNA,
DEFINITION complete cds.
ACCESSION AF034133
VERSION AF034133.1 GI:2921337
KEYWORDS
SOURCE
ORGANISM
upland cotton.
Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 1033)
Loguerico, L.L., Zhang, J.Q. and Wilkins, T.A.
Differential regulation of six novel MYB-domain genes defines two
distinct expression patterns in allotetraploid cotton (Gossypium
hirsutum L.)
Mol. Gen. Genet. 261 (4-5), 660-671 (1999)
JOURNAL
MEDLINE 99320826
PUBMED 10394903
REFERENCE
AUTHORS Loguerico, L.L., Zhang, J. and Wilkins, T.A.
TITLE Direct Submission
JOURNAL Submitted (13-NOV-1997) Agronomy & Range Science, University of
California, One Shields Ave., Davis, CA 95616-8515, USA
FEATURES
source
1.1033
/organism="Gossypium hirsutum"
/cultivar="Acala SJ-2"
/db_xref="taxon:3635"
/tissue_type="ovule"
/dev_stage="3 days pre-anthesis"
1.1033
/gene="MYB5"
/note="MYB-domain gene N"
106.987
/gene="MYB5"
/function="putative MYB-like transcription factor"
/note="similar to MYB A encoded by Genbank Accession
Number I04497"
/codon_start=1
/product="MYB-like DNA-binding domain protein"
/protein_id="AAC04719.1"
/translation="MFTREYSYLKLEVKRWRIRKQPTTEEDSMIRAYVINIGEGR
WNAVARLSGLRFTGKSLRLWLMYMRPEIKRNLISLEQLILIELSRWGNRWSIAQ
HLPGRDNEIKNYKTRVVOAKOLNCDVNSKQFKDAMRYVWIPRLVERIRASSESPS
SOPSTTNTYNDIRISNISSSOMSYANASVQVDPSSLPELGGTSSDIDTQVSYS
DLDCCYPOSTSYLHKIGLEKEGAATMGREDFQATEHSHNGWLVGGESEMDIYV
NEENWFLQOQLHDDGI"
repeat_region
163.318
/note="putative MYB DNA-binding domain repeat R2"
repeat_region
319.474
/note="putative MYB DNA-binding domain repeat R3"
BASE COUNT 308 a 223 c 253 g 249 t
ORIGIN
Query Match 16.9%; Score 146.2; DB 8; Length 1033;

Best Local Similarity 66.1%; Pred. No. 1.6e-21;
Matches 211; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 72 TGAGAAAGGCGCTTGACATGGAAGAGACTGATCTTGATGATATATTCGAATC 131
DB 164 TAAGAAAGGCTCCATGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 223
QY 132 ATGGGGAAGGTGTTGGAACTCTTTGGCCAAAGCTGCTGCTCAAAAGTAAAGCA 191
DB 224 ACGGGAAGGTGCTGGAACGCTGTGCTTATGAGATTAAGAAACCGGTAAA 283
QY 192 GTWCCGCTAAGGTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 251
DB 284 GCTGCACATTAAGATGCTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 343
QY 252 CCGAGAACAACTTTGATTTATGAGACTCCACGCAAAAGTGGGAAACAGGTGCAAAA 311
DB 344 TCGAAGAGCAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 403
QY 312 TTGCAAGCAGCTGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 371
DB 404 TTGCAAGCAGCTGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 463
QY 372 TCCAGAGCAGCATCAAGCA 390
DB 464 TCCAGAGCAGCAGCAAGCA 482

Search completed: February 18, 2003, 08:50:52
Job time : 2184 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 08:10:00 ; Search time 48 Seconds
(without alignments)
5513.793 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863
Sequence: 1 gcacgagctctatcacacac.....aaaaaaaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/PTCUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112.4	13.0	2220	4	US-08-997-251-1 Sequence 1, Appli
2	110.8	12.8	2352	4	US-08-997-251-3 Sequence 3, Appli
3	108.2	12.5	1344	2	US-08-722-626B-1 Sequence 1, Appli
4	87	10.1	3602	4	US-09-402-923-1 Sequence 1, Appli
5	87	10.1	6775	4	US-09-402-923-4 Sequence 4, Appli
6	83.8	9.7	4880	4	US-09-402-923-5 Sequence 5, Appli
7	78.8	9.1	1035	5	PCT-US93-06251-90 Sequence 90, Appli
8	78.8	9.1	3225	5	US-08-306-691B-45 Sequence 45, Appli
9	78.8	9.1	3225	5	PCT-US93-06251-91 Sequence 91, Appli
10	78.8	9.1	3230	5	PCT-US93-06251-89 Sequence 89, Appli
11	72.4	8.4	1897	4	US-09-167-322-10 Sequence 10, Appli
12	60	7.0	2638	1	US-08-306-691B-46 Sequence 46, Appli
13	52.8	6.1	7218	4	US-08-232-463-14 Sequence 14, Appli
14	51.6	6.0	3573	4	US-09-353-585-4 Sequence 4, Appli
15	51.2	5.9	1007	4	US-09-171-156A-61 Sequence 61, Appli
16	51.2	5.9	1007	4	US-09-171-156A-63 Sequence 63, Appli
17	51	5.9	1641	1	US-08-300-903A-8 Sequence 8, Appli
18	51	5.9	4059	2	US-08-485-139-1 Sequence 1, Appli
19	51	5.9	4059	3	US-08-485-139-1 Sequence 1, Appli
20	51	5.9	4824	2	US-08-485-139-5 Sequence 5, Appli
21	51	5.9	4824	3	US-08-485-139-5 Sequence 5, Appli
22	50.4	5.8	2007	3	US-08-747-221B-36 Sequence 36, Appli
23	50.4	5.8	2007	3	US-08-747-221B-38 Sequence 38, Appli
24	50.4	5.8	2007	4	US-09-005-051-36 Sequence 36, Appli
25	50.4	5.8	2007	4	US-09-005-051-38 Sequence 38, Appli
26	49.4	5.7	3138	1	US-07-867-106-4 Sequence 4, Appli
27	48.8	5.7	1582	3	US-08-545-196B-10 Sequence 10, Appli

28	48.8	5.7	1582	3	US-08-545-196B-12	Sequence 12, Appli
29	48	5.6	1577	4	US-08-821-994-59	Sequence 59, Appli
30	47.8	5.5	1313	4	US-09-149-476-112	Sequence 112, App
31	47.8	5.5	1474	4	US-08-821-994-64	Sequence 64, Appli
32	47.6	5.5	785	3	US-09-008-979A-6	Sequence 6, Appli
33	47.6	5.5	785	4	US-09-460-618-6	Sequence 6, Appli
34	47.6	5.5	785	4	US-09-310-235B-6	Sequence 6, Appli
35	47.4	5.5	1441	4	US-08-821-994-63	Sequence 63, Appli
36	47.4	5.5	5852	1	US-07-867-106-2	Sequence 2, Appli
37	47	5.4	370	4	US-09-376-113-1	Sequence 1, Appli
38	45.8	5.3	1100	3	US-09-248-335-33	Sequence 53, Appli
39	45.8	5.3	1307	4	US-09-250-609-3	Sequence 3, Appli
40	45.8	5.3	1753	4	US-09-149-476-56	Sequence 56, Appli
41	45.6	5.3	2908	4	US-09-930-181-1	Sequence 1, Appli
42	45.4	5.3	1781	4	US-09-499-302A-1	Sequence 1, Appli
43	45.2	5.2	760	1	US-08-276-452A-49	Sequence 49, Appli
44	45.2	5.2	760	2	US-08-798-744-49	Sequence 49, Appli
45	45.2	5.2	1889	3	US-09-187-050-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-997-251-1
; Sequence 1, Application US/08997251
; Patent No. 6271440
; GENERAL INFORMATION:
; APPLICANT: GUTBER, FRANZ J.
; APPLICANT: JACOBSEN, JOHN V.
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Boulevard
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,251
; FILING DATE: 23-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO AU96/00383
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN6470/95
; FILING DATE: 09-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN3779/95
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 110-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORGANISM: Hordeum vulgare

APPLICATION NUMBER: US/08/7122,626B
 FILING DATE: 27-SEP-1996
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Pat Hagan
 REGISTRATION NUMBER: 27,643
 REFERENCE/DOCKET NUMBER: 97-0010
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215 563-4100
 TELEFAX: 215 563-4044
 TELEX:
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1344 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 148..981
 OTHER INFORMATION:
 US-08-7122-626B-1

ADDRESSSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
 STREET: Suite 1800 Two Penn Center Plaza
 CITY: Philadelphia
 STATE: PA
 COUNTRY: U.S.A.
 ZIP: 19102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/402.929
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US98/06896
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Monaco, Daniel A.
 REGISTRATION NUMBER: 30,480
 REFERENCE/POCKET NUMBER: 6056-214 PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-8383
 TELEFAX: (215) 568-5549
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3602 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-402-929-1

```

:
: APPLICANT: Hatton, Kimi
:
: TITLE OF INVENTION: A-myd NULL MUTANT TRANSGENIC ANIMALS AND
:
: NUMBER OF SEQUENCES: 7
:
: CORRESPONDENCE ADDRESSES:
:   ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
:   STREET: Suite 1800 Two Penn Center Plaza
:   CITY: Philadelphia
:   STATE: PA
:   COUNTRY: U.S.A.
:   ZIP: 19102
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/402,929
:   FILING DATE:
:
: CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: PCT/US98/06896
:   FILING DATE:
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Monaco, Daniel A.
:     REGISTRATION NUMBER: 30,480
:     REFERENCE/DOCKET NUMBER: 6056-214 PC
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (215) 568-8383
:       TELEFAX: (215) 568-5549
:     INFORMATION FOR SEQ ID NO: 4:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 6775 base pairs
:         TYPE: nucleic acid
:         STRANDEDNESS: single
:         TOPOLOGY: linear
:
: US-09-402-929-4

```

```

Query Match          10.1%; Score 87; DB 4; Length 6775;
Best Local Similarity 54.2%; Pred. No. 1.3e-11;
Matches 199; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 39 AACAACTGTAGACGCTCAGATCCTGAAAGTGAAGAAAGGCGCTTGACAAATGGAG 98
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3726 AACATCGATGCGAAGAGTTTAAATCCAGAAATTGATTAAGGCTTGACAAAGAG 3785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 99 AAGACTGTATGATGAACATAATTCGAATCATGGGGAAGGTGTTGAACTCTTG 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3786 AAGATCAGAGGCTTATTTAGATTAGTTTCAGAAATATGCGCCAAAAGGTGCTTTAATTG 3845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 159 CCAAGCTGCTGCTCTCAACAGTAACGAAAGAGTTCGCCGCTAAGTGGCTAAATTAC 218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3846 CAAACATTTTAAAGGAAG--ATAGGCAAGCAGTGCAGAAAGATGGCACAATCACC 3902
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 219 TCCGTCCTGATGTAGAAAGAGGAATATTACACCCGAGAACCACTTTGTTATGGAGC 278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3903 TGAACCTGGAAGTGAAGAGCTTCCTCGACAGAAAGAAAGACAGCATCTATATATGAG 3962
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 TCCAGCGAAAGTGGGAAACAGGTGTCCAAAATTTGCCAACATCTACTGGAAGAGCTG 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3963 CACACAAAGCGCTGGGAAACCGTTGGGCGGAGATTGCTAGTTACTTCTGGAAGAGCTG 4022,
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 ATAATGAGATCAAGAACTATTGGAGAGCAAGATCCAGAAACACATCAAGCAAGCTGAG 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4023 ATAATTTCTATCAAAAATCATTTGGAATTTCTACATGCCAAGAAAGTGAACAGAGGCT 4082
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 399 ACTTCA 405
    |||||
DB 4083 ATTACA 4089
    |||||

```

RESULT 6

```

US-09-402-929-5
: Sequence 5; Application US/09402929
: Patent No. 6410825
:
: GENERAL INFORMATION:
:   APPLICANT: Temple University - Of the Commonwealth System of Higher Education
:   APPLICANT: Toscani, Antonio
:   APPLICANT: Hatton, Kimi
:   APPLICANT: Reddy, E. P.
:   TITLE OF INVENTION: A-myd NULL MUTANT TRANSGENIC ANIMALS AND
:   NUMBER OF SEQUENCES: 7
:   CORRESPONDENCE ADDRESSES:
:     ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
:     STREET: Suite 1800 Two Penn Center Plaza
:     CITY: Philadelphia
:     STATE: PA
:     COUNTRY: U.S.A.
:     ZIP: 19102
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/09/402,929
:   FILING DATE:
:
: CLASSIFICATION:
:   PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: PCT/US98/06896
:   FILING DATE:
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Monaco, Daniel A.
:     REGISTRATION NUMBER: 30,480
:     REFERENCE/DOCKET NUMBER: 6056-214 PC
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (215) 568-8383
:       TELEFAX: (215) 568-5549
:     INFORMATION FOR SEQ ID NO: 5:
:       SEQUENCE CHARACTERISTICS:
:         LENGTH: 4880 base pairs
:         TYPE: nucleic acid
:         STRANDEDNESS: single
:         TOPOLOGY: linear
:
: US-09-402-929-5

```

```

Query Match          9.7%; Score 83.8; DB 4; Length 4880;
Best Local Similarity 53.7%; Pred. No. 6.5e-11;
Matches 197; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 39 AACAACTGTAGACGCTCAGATCCTGAAAGTGAAGAAAGGCGCTTGACAAATGGAG 98
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 325 AGCATGATGCGAAGAGTTTAAATCCAGAAATTGATTAAGGCTTGACAAAGAG 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 99 AAGACTGTATGATGAACATAATTCGAATCATGGGGAAGGTGTTGAACTCTTG 158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 AAGATCAGAGGCTTATTTAGATTAGTTTCAGAAATATGCGCCAAAAGGTGCTTTAATTG 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 159 CCAAGCTGCTGCTCTCAACAGTAACGAAAGAGTTCGCCGCTAAGTGGCTAAATTAC 218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 445 CAAACATTTTAAAGGAAG--ATAGGCAAGCAGTGTGAGAAAGATGGCATAATCACC 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 219 TCCGTCCTGATGTAGAAAGAGGAATATTACACCCGAGAACCACTTTGTTATGGAGC 278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 502 TGAATCCTGAGTGAAGAAATCTTCCTGACAGAAAGAGAGACAGATCATCTATGAG 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 TCCAGCGAAAGTGGGAAACAGGTGTCCAAAATTTGCCAACATCTACTGGAAGAGCTG 338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 562 CACATAAGCGGTTGGGAAATCGTTGGGCGAAGAAATTTGCCAAACTACTCCAGAGAGGCTG 621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 339 ATAATGAGATCAAGAACTATTGGAGAGCAAGATCCAGAAACACATCAAGCAAGCTGAGA 398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 622 ATAATTTCTATCAAAAATCATTTGGAATTTCTACTATGGAAGAAAGTGAACAGAGGCT 681
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

OY 399 ACTTTCA 405
| | | | |
Db 682 ATTACCA 688

RESULT 7

PCT-US93-06251-90
; Sequence 90, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: PCT/US93/06251
; APPLICATION NUMBER: 19930630
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digilio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1035 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US93-06251-90

Query Match 9.1%; Score 78.8; DB 5; Length 1035;
Best Local Similarity 52.7%; Pred. No. 6.8e-10;

Matches 195; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

OY 39 AACAACTGTTAAGACGCTTCAAGATCTGTAAGTGAAGAAAGGCGCTTGACAATGGAAG 98
| | | | |
Db 98 AGACCGATGGCAGAAAGTACTAAACCTGACCTCATCAAGAGGCTCTGGACCAAGAAG 157
| | | | |
OY 99 AAGACTTGATCTGATGAATATATATGCAATATATGAGGAGAGTGTGGAACTCTTTGG 158
| | | | |
Db 158 AAGATCAGAGAGTATGAGCTTGTACAGAAATACGCTCCGAACGTTGCTGTATTGG 217
| | | | |
OY 159 CCAAGCTGCTGCTCAACAGTACGGAAGAGTTGCCGCTAAGTGGCTAAATTACC 218
| | | | |
Db 218 CCAAGACCTTAAGGGA--GAATTGGAAGAAATGTAAGGAGAGTGTGCATTAACCACT 274
| | | | |
OY 219 TCCGTCCTGATTTAAGAGGGAATATTACACCCGAGAACCACTTTTGAATTANGGAGC 278
| | | | |
Db 275 TGAATCCAGAAAGTAAAGAAACCTCTCGACAGAAAGAGACAGAAATATTATTACCAAG 334
| | | | |
OY 279 TCCACCAAGAGTGGGAAACAGGTGTCACAAATTTGCCAAGATCTACTGGAAGACAGC 338
| | | | |
Db 335 CACACAGAGAGCTGGGAAACAGATGGGACAGAAATTCGCAAGAGCTACTGCTGGACGAGCAG 394
| | | | |
OY 339 ATATAGATCAAGAACTATTGGAGAGCAAGATCCAGAGACACATCAAGCAAGCTGACA 398
| | | | |
Db 395 ATATATGATCAAGAACCACTGGAATTTTCAATATGCGTGGAAGAGTTCGAAACGGAAGTT 454
| | | | |

OY 399 ACTTTCAGCA 408
| | | | |
Db 455 ATCTGCAGCA 464

RESULT 8

US-08-306-691B-45
; Sequence 45, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavourga & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA: US/08/306,691B
; APPLICATION NUMBER: 514
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3225 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-306-691B-45

Query Match 9.1%; Score 78.8; DB 1; Length 3225;
Best Local Similarity 52.7%; Pred. No. 8.8e-10;

Matches 195; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

OY 39 AACAACTGTTAAGACGCTTCAAGATCTGTAAGTGAAGAAAGGCGCTTGACAATGGAAG 98
| | | | |
Db 349 AGACCGATGGCAGAAAGTACTAAACCTGACCTCATCAAGAGGCTCTGGACCAAGAAG 408
| | | | |
OY 99 AAGACTTGATCTGATGAATATATGCAATATATGAGGAGAGTGTGGAACTCTTTGG 158
| | | | |
Db 409 AAGATCAGAGAGTATGAGCTTGTACAGAAATACGCTCCGAACGTTGCTGTATTGG 468
| | | | |
OY 159 CCAAGCTGCTGCTCAACAGTACGGAAGAGTTGCCGCTAAGTGGCTAAATTACC 218
| | | | |
Db 469 CCAAGCACTTAAGGGA--GAATTGGAAGAAATGTAAGGAGAGTGTGCATTAACCACT 525
| | | | |
OY 219 TCCGTCCTGATTTAAGAGGGAATATTACACCCGAGAGCAACTTTTGAATTANGGAGC 278
| | | | |
Db 526 TGAATCCAGAAAGTAAAGAAACCTCTCGACAGAAAGAGAGCAAGATATTATTACCAAG 585
| | | | |
OY 279 TCCACCAAGAGTGGGAAACAGGTGTCACAAATTTGCCAAGATCTACTGGAAGAGCAGC 338
| | | | |

Db	586	CACACAAAGAGACTGGGGGAACGATGATGGGACAGAAATCGGAAAAGCTACTCCTCGACAGCAACTG	645
Qy	339	ATATATGATGATAAACAATGATGGAGGACCAAGATCCGAAAGACCATGACCAAGAGCTGAGA	398
Db	646	ATATATGCTATCAAGAACCACTGGAAATCTCTACAAATGCGCTCGGAAGGTGGAACAGGAAGTT	705
Qy	399	ACTTTCAGCA	408
Db	706	ATCTGCAGGA	715

RESULT 9
 PCT-US93-06251-91
 Sequence 91, Application PC/WUS9306251
 GENERAL INFORMATION:
 APPLICANT: WICKSTROM, Eric and Rife, Jason P.
 TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
 TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: NY
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/06251
 FILING DATE: 19930630
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: DIGILIO, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8586
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3225 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PCT-US93-06251-91

Accession	Sequence	Position
Db	586 CACACAGAAGACTGGGGAAACAGATGGGGAGAAATCGCAAGCTACTGCTGGAGCAACTG	645
OY	339 ATATGAGATCAAGAACTATTGGAGAGACAGATCCAGAACACATCAAGCAAGCTAGA	338
Db	646 ATAAAGCTGATCAAGAACCACTGGAAATCTACAAATGCCCTCGGAAGGTGAACAGGAAGTT	705
OY	399 ACTTTCAGCA 408	
Db	706 ATCTGCAGGA 715	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tomato
; STRAIN: Cf2
; SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-09-353-585-4

Query Match
Best Local Similarity 51.8%; Pred. No. 0.0021;
Matches 114; Conservative 1; Mismatches 105; Indels 0; Gaps 0;

Qy 644 GATTAATATATATCAAGATTAACCTAAGTTTGAGTTCATACGCTGGAATGCTCT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3352 GATTGATTTCAAGACTTCGACTTCAGAGCCAGAAATAGAAACGCTGCTAAG 3411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 704 GATTAACATATATTTGGCTTGTATATAGTAGTGGATGTTTGGCTAC 763
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3412 GATTTGCTCTCTCGTGGTGCAGCTTATGATGTGATAGATTGATTATTAAGC 3471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 764 CATATATAGCTATGCTGTATATATACAGATTTATATTAACTATATCGCATGCT 823
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3472 TTTTCTTCACTTGAGAAATGTAATATTTATGAATTTGATGATTAACAATTAAGTGTGT 3531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 824 TTATATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3532 TTATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-171-156A-61
; Sequence 61, Application US/09171156A
; Patent No. 6368846
; GENERAL INFORMATION:
; APPLICANT: Hunter, Shirley Wu
; Sm, Gek-kee
; Weber, Eric R.
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
; APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHERIDAN ROSS P.C.
; STREET: 1560 BROADWAY, SUITE 1200
; CITY: DENVER
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,156A
; FILING DATE: 04-Mar-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```

;
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..465
; SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-09-171-156A-61

Query Match
Best Local Similarity 55.9%; Pred. No. 0.002;
Matches 95; Conservative 1; Mismatches 74; Indels 0; Gaps 0;

Qy 692 TGAATGCTCTTGATTAACATATTTAGGCTTGTATATAGAGTGGATGTT 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838 TGGGATGTTATATATATATAAAGTAGTGATTTATTTCTTAAGAAAGTTGCANAA 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 752 GCTTTGCCATCAATTAATAGCTATGCTGTAATATATACAGATTTATATTAAC 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898 TGTTCATATCAATATACCGAATATTTCAATATATATATGTATGATTAATCGC 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 812 TATCTGATGCTTTATATATATAAAAAAAAAAAAAAAAAAAAAA 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 958 CATCTGATGCTTAATATATAAAAAAAAAAAAAAAAAAAAAA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 18, 2003, 09:20:41
Job time : 71 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:07:50 ; Search time 218 Seconds

(without alignments)
8915.021 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863
Sequence: 1 gcacgagcctctacacac.....aaaaaaaaaaaaaaaa 863

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq-101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	248.4	28.8	862	21	AAC57336
2	239.4	27.7	910	21	AAC40587
3	236.2	27.4	908	21	AAC48506
4	223	25.8	907	21	AAC57337
5	208	24.1	621	21	AAC41714
6	208	24.1	1127	21	AAC37953
7	175	20.3	626	22	AAH87724
8	170.6	19.8	1137	22	AAD05784
9	169.4	19.6	837	20	AAK25572

10	169.4	19.6	981	24	ABK65252	Arabidopsis cDNA e
11	165.2	19.1	972	21	AAC56197	Eucaalyptus grandis
12	163.4	18.9	723	24	ABK65364	Arabidopsis cDNA e
13	161.6	18.7	389	21	AAC56152	Eucaalyptus grandis
14	161.6	18.7	417	21	AAC56774	Eucaalyptus grandis
15	152.4	17.7	524	21	AAC57194	Eucaalyptus grandis
16	143.2	16.6	959	21	AAC44894	Arabidopsis thalia
17	132.8	15.4	516	21	AAC56352	Pinus radiata tran
18	132.8	15.4	516	21	AAC56457	Pinus radiata tran
19	131.4	15.2	542	21	AAC56434	Pinus radiata tran
20	130.8	15.2	839	21	AAC42479	Pinus radiata tran
21	130.8	15.2	918	22	AAD06466	Arabidopsis thalia
22	130.6	15.1	389	21	AAC56837	Arabidopsis thalia
23	129.6	15.0	984	21	AAC44207	Pinus radiata tran
24	128.6	14.9	1825	21	AAC44207	Arabidopsis thalia
25	128.4	14.8	1329	21	AAC49321	Arabidopsis thalia
26	128	14.8	1329	21	AAC49321	Arabidopsis thalia
27	127	14.7	1205	21	AAC47877	Arabidopsis thalia
28	127	14.7	1205	21	AAC47877	Arabidopsis thalia
29	126.8	14.7	1332	24	ABK65304	Arabidopsis thalia
30	125.4	14.5	734	21	AAC54477	Arabidopsis thalia
31	125.4	14.5	736	21	AAC52742	Arabidopsis thalia
32	125.2	14.5	612	24	ABK65317	Arabidopsis cDNA e
33	124.4	14.4	1152	21	AAC46330	Arabidopsis thalia
34	124.4	14.4	1153	21	AAC39525	Arabidopsis thalia
35	123.2	14.3	373	21	AAC57227	Eucaalyptus grandis
36	123	14.3	378	21	AAC56151	Eucaalyptus grandis
37	122.8	14.2	639	21	AAC51202	Arabidopsis thalia
38	122.8	14.2	916	22	AAD05774	Arabidopsis thalia
39	122.6	14.2	825	21	AAC42974	Arabidopsis thalia
40	122.6	14.2	825	22	AAD06465	Arabidopsis thalia
41	121.6	14.1	1151	22	AAF90595	Cotton transcript
42	121.6	14.1	1300	24	ABK65274	Arabidopsis thalia
43	120.8	14.0	774	21	AAC42662	Arabidopsis thalia
44	120.8	14.0	1044	24	ABN98399	Arabidopsis thalia
45	120.6	14.0	1678	21	AAC57334	Eucaalyptus grandis

ALIGNMENTS

RESULT 1
ID AAC57336 standard; DNA: 862 BP.

AC AAC57336;

DT 25-JAN-2001 (first entry)

DE Eucaalyptus grandis transcription factor DNA sequence #773.

KW Plant; transcription factor; gene expression; eucaalyptus; pine; acacia;

KW Poplar; sweetgum; teak; mahogany; bzip; G-box binding factor;

KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;

KW homeodomain zipper; LIM domain; AP2; ERBBS; zinc finger domain;

OS type 2 Cys2His2; CCNAT box element; MYB; ss.

XX Eucaalyptus grandis.

PN WO200053724-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06112.

PR 11-MAR-1999; 99US-0266513.

PR 18-AUG-1999; 99US-0149485.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

PI Wood M, McGrath A, Shenk MA, Glenn M;

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.7%; Score 239.4; DB 21; Length 910;
Best Local Similarity 72.4%; Pred. No. 1e-38;
Matches 325; Conservative 0; Mismatches 121; Indels 3; Gaps 1;
QY 56 TCTCAAGATCCGAGAGTGAAGAAAGGCGCTTGACATGAGAGAGACTGATCTTGATG 115
DB 143 TCAGGAGATGCAGAGGTAGAAAGGCCATGACGATGAGAAAGATTATCTCTATC 202
QY 116 AACATATTGCAAAATCATGCGGAAGGTGTTGGAACTCTTGGCCAAAGCTCGTGCCTC 175
DB 203 AATTATTCGCCAATCATCATGTTGAGGTGTTGAACTCTTCGCCAAATCTGCAGAGACTA 262
QY 176 AAACGTAAACGGAAGAGTTGCCGGCTAAGTGGCTAAATTAACCTCCGTCGATGTAGA 235
DB 263 AAACGCACCGGGAAGAAAGTTGCCGGCTCGGCTGAGTGAACCTACCTCCGACCTGATGCGA 322
QY 236 AGAGGAAATTTACACCCCGAGGAACAACCTTTGATATGAGCTCACGCAAGTGGGA 295
DB 323 CGGGAAATATCACACCGAAGAACAGCTCACATCATGAACTTACGCAAAATGGGGA 382
QY 296 AACAGTGTCCAAAAATTGCCACAGCATCTACCTGAAGAGACTGATTAATGAGATCAAGAC 355
DB 383 AATAGTGTGCATAAAATTTGCAAAAGCATTTACAGAGGAAGGACCGACATGAGATTAAGAC 442

Query Match	27.4%	Score 236.2	DB 21	Length 908
Best Local Similarity	71.9%	Pred. No. 4.5e-38		
Matches 323; Conservative	0	Mismatches 123;	Indels 3	Gaps 1;

producing plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide

PS Claim 1; Pages 649-650; 747pp; English.

XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper,
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and PEBBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.

XX Sequence 907 BP; 280 A; 210 C; 212 G; 205 T; 0 other:

Query Match 25.8%; Score 223; DB 21; Length 907;

Best Local Similarity 77.2%; Pred. No. 1.9e-35; Mismatches 271; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 36 AAAAAAAGCTGAAGCGCTCAAGATCTGTAAGTGAAGAAAGGCGCTTGACATGG 95
DB 8 ACAAAGAAACCGTGTACCGCAGCAGATCTCTCAAGTGAAGAAAGCAGCTCTTG 67
QY 96 AAGAAGACTGTATCTGTATGAAGAACTATTTGCAAAATCATGGGGAAGGTGTTGAACTCTT 155
DB 68 AAGAGATCTGTATCTGTATGATGATCAATAGCAACCATGGCGAAGGTATGAACTCTC 127
QY 156 TGGGCAAGCGCTGCGTCAAGCTTAAGGAAAGAGTTCGCGCTTAAGGCGCTTAATT 215
DB 128 TAGCTAAAGCTGCTGCTCTTCAACGCAACGAAAGAGCTCGCGCTCGATGCTGAATT 167
QY 216 ACCTCCGCTCTGATGTATGAAGAGGAATTTACACCCGAGAGAAACAATTGATTATGG 275
DB 188 ACCTCCGAGCCGATGTCGAGAGGTATATCACTCCGAGAAAGACACTGTGATCATAC 247
QY 276 AGCTCCAGCAAGTGGGGAAGAGGTGTCCAAAATTTGCCAATCATCTACCTGGAAGA 335
DB 248 ACCTGCAATCATGTGGGGAAGAGGTGTCGAAATCGGAATCTTCCGGGGAAGA 307
QY 336 CTGATATGATCAAGAACTATTTGAGAGCAAGATCCAGAACATCA 386
DB 308 CCGACAAATGAATTAAGAACTACTGAGAGCAACCAATCCAAAGCACATTA 358

RESULT 5

ID AAC41714 standard; DNA; 621 BP.

AC AAC41714;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 32868.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SHP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123160.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132487.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144332.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.1%; Score 208; DB 21; Length 621;
Best Local Similarity 77.1%; Pred. No. 1.8e-32;
Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 60 AAGATCCATAAGTGAAGAAAGGCCCTTGACACATGGAAGAGACTTGATCTTGATGAAC 119
DB 62 AAGAGGAACAGTGAAGAAAGGACCTTGACATGATGAAAGAAATTCCTTTAAAT 121
QY 120 AATTGCAATCATGAGGAAGGTGTTGGAACTTTTGACCAAGCTGCTGCTCAAG 179
DB 122 ACATCCTTATCATGATGTAAGGTCTTTGAACTCTGTGCCAAAGCCTTGCTTAAG 181
QY 180 GTACGGAAGAAAGTGTGCGTCCGTCGATGATCTTCACTCGACACAGATGCGCGAG 241
DB 182 GTACTGGAAGAAAGTGTGCGTCCGTCGATGATCTTCACTCGACACAGATGCGCGAG 241
QY 240 GGAATTTTACACCCGAGAACACTTTGATATGAGCTCCAGCAAGTGGGAACA 299
DB 242 GGAACATTAACCGCAGAAAGACGCTTTGATCATTCAGCTTCATGCTTAAGCTTGA 301
QY 300 GGTGTCGCAAAATGCGACAGCATCTACCTGGAAGACTGATGAATCAAGAACTATT 359
DB 302 GGTGTCGCAAGATGCGAAGCATCTTCGCGGAAGAACGACACAGATTAAGAACTTCT 361
QY 360 GGAGGACAAAGATCCAGACACATCAA 387
DB 362 GGAGGACAAAGATTCAGACACATGAA 389

RESULT 6
AAC37953
ID AAC37953 standard; DNA: 1127 BP.
XX AAC37953;
AC
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 19258.
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX

OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX	06-SEP-2000.	
PD		
XX		
PR	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	PR 30-JUN-1999; 99US-0141287.
PR	05-MAR-1999; 99US-0123180.	PR 01-JUL-1999; 99US-0141842.
PR	09-MAR-1999; 99US-0123548.	PR 01-JUL-1999; 99US-0142154.
PR	23-MAR-1999; 99US-0125788.	PR 02-JUL-1999; 99US-0142055.
PR	25-MAR-1999; 99US-0126264.	PR 06-JUL-1999; 99US-0142390.
PR	29-MAR-1999; 99US-0126785.	PR 08-JUL-1999; 99US-0142803.
PR	01-APR-1999; 99US-0127462.	PR 09-JUL-1999; 99US-0142920.
PR	06-APR-1999; 99US-0128234.	PR 12-JUL-1999; 99US-0142977.
PR	08-APR-1999; 99US-0128714.	PR 13-JUL-1999; 99US-0143542.
PR	16-APR-1999; 99US-0129845.	PR 14-JUL-1999; 99US-0143624.
PR	19-APR-1999; 99US-0130077.	PR 15-JUL-1999; 99US-0144005.
PR	21-APR-1999; 99US-0130449.	PR 16-JUL-1999; 99US-0144085.
PR	23-APR-1999; 99US-0130510.	PR 16-JUL-1999; 99US-0144086.
PR	23-APR-1999; 99US-0130891.	PR 19-JUL-1999; 99US-0144325.
PR	28-APR-1999; 99US-0131449.	PR 19-JUL-1999; 99US-0144331.
PR	30-APR-1999; 99US-0132048.	PR 19-JUL-1999; 99US-0144332.
PR	30-APR-1999; 99US-0132407.	PR 19-JUL-1999; 99US-0144333.
PR	04-MAY-1999; 99US-0132484.	PR 19-JUL-1999; 99US-0144334.
PR	05-MAY-1999; 99US-0132485.	PR 19-JUL-1999; 99US-0144335.
PR	06-MAY-1999; 99US-0132486.	PR 20-JUL-1999; 99US-0144352.
PR	06-MAY-1999; 99US-0132487.	PR 20-JUL-1999; 99US-0144632.
PR	07-MAY-1999; 99US-0132863.	PR 21-JUL-1999; 99US-0144814.
PR	11-MAY-1999; 99US-0134256.	PR 21-JUL-1999; 99US-0145086.
PR	14-MAY-1999; 99US-0134218.	PR 21-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134219.	PR 22-JUL-1999; 99US-0145087.
PR	14-MAY-1999; 99US-0134221.	PR 22-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134370.	PR 22-JUL-1999; 99US-0145089.
PR	18-MAY-1999; 99US-0134768.	PR 22-JUL-1999; 99US-0145192.
PR	19-MAY-1999; 99US-0134941.	PR 23-JUL-1999; 99US-0145145.
PR	20-MAY-1999; 99US-0135124.	PR 23-JUL-1999; 99US-0145218.
PR	21-MAY-1999; 99US-0135353.	PR 26-JUL-1999; 99US-0145276.
PR	24-MAY-1999; 99US-0135629.	PR 27-JUL-1999; 99US-0145913.
PR	25-MAY-1999; 99US-0136021.	PR 27-JUL-1999; 99US-0145919.
PR	27-MAY-1999; 99US-0136382.	PR 28-JUL-1999; 99US-0145951.
PR	28-MAY-1999; 99US-0136782.	PR 02-AUG-1999; 99US-0146386.
PR	01-JUN-1999; 99US-0137222.	PR 02-AUG-1999; 99US-0146388.
PR	03-JUN-1999; 99US-0137528.	PR 02-AUG-1999; 99US-0146389.
PR	04-JUN-1999; 99US-0137502.	PR 03-AUG-1999; 99US-0147038.
PR	07-JUN-1999; 99US-0137724.	PR 04-AUG-1999; 99US-0147204.
PR	08-JUN-1999; 99US-0138094.	PR 04-AUG-1999; 99US-0147302.
PR	10-JUN-1999; 99US-0138540.	PR 05-AUG-1999; 99US-0147192.
PR	10-JUN-1999; 99US-0138847.	PR 05-AUG-1999; 99US-0147260.
PR	14-JUN-1999; 99US-0139119.	PR 06-AUG-1999; 99US-0147303.
PR	16-JUN-1999; 99US-0139452.	PR 09-AUG-1999; 99US-0147493.
PR	16-JUN-1999; 99US-0139453.	PR 10-AUG-1999; 99US-0148171.
PR	17-JUN-1999; 99US-0139492.	PR 11-AUG-1999; 99US-0148319.
PR	18-JUN-1999; 99US-0139454.	PR 12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139455.	PR 13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139456.	PR 13-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139457.	PR 16-AUG-1999; 99US-0149368.
PR	18-JUN-1999; 99US-0139458.	PR 17-AUG-1999; 99US-0149175.
PR	18-JUN-1999; 99US-0139459.	PR 18-AUG-1999; 99US-0149426.
PR	18-JUN-1999; 99US-0139460.	PR 20-AUG-1999; 99US-0149723.
PR	18-JUN-1999; 99US-0139461.	PR 20-AUG-1999; 99US-0149929.
PR	18-JUN-1999; 99US-0139462.	PR 23-AUG-1999; 99US-0149902.
PR	18-JUN-1999; 99US-0139463.	PR 23-AUG-1999; 99US-0149930.
PR	18-JUN-1999; 99US-0139750.	PR 25-AUG-1999; 99US-0150566.
PR	18-JUN-1999; 99US-0139753.	PR 26-AUG-1999; 99US-0150884.
PR	21-JUN-1999; 99US-0139817.	PR 27-AUG-1999; 99US-0151065.
PR	22-JUN-1999; 99US-0139839.	PR 27-AUG-1999; 99US-0151066.
PR	23-JUN-1999; 99US-0140353.	PR 30-AUG-1999; 99US-0151080.
PR	23-JUN-1999; 99US-0140354.	PR 31-AUG-1999; 99US-0151303.
PR	24-JUN-1999; 99US-0140655.	PR 01-SEP-1999; 99US-0151338.
PR	28-JUN-1999; 99US-0140823.	PR 07-SEP-1999; 99US-0152363.
PR	29-JUN-1999; 99US-0140991.	PR 10-SEP-1999; 99US-0153070.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.1%; Score 208; DB 21; Length 1127;
Best Local Similarity 77.1%; Pred. No. 1.9e-32;
Matches 253; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 60 AAGATCCTGAAGTGAAGAAAGGCGCTTGACATGGAAGAAAGACTTGATCTTGATGAACT 119
DB 335 AAGAGGAGACGTGAGAAAGAGACCTTGACATGGAAGAAAGATTCATCCTCTTTAATT 394
QY 120 ATATTGCAATCATGGGGAAGGTGTTGGAACCTTTGGCCAAAGCTGTCGTCGAAC 179
DB 395 ACATCTTAATCATGATGGGAAGGTCTTTGGAACTCTGTCGCCAAAGCCTGTCGTTAAAC 454
QY 180 GTAACGGAAGAGTGGCCGGCTAAGTGCGCTAAATTAACCTCGCTCGATGTTAGAAAG 239
DB 455 GTACTGCAAAAAGTTGCGGCTCGGCTGCGACTATCTCGACGACGATGTGCGGCGAG 514
QY 240 GGATATTACACCGGAGGAACAATTTGATATGAGAGCTCCAGCAAGGGGGAACA 299
DB 515 GGAACATTAACCAAGAAAGAACGCTTTTGATCATTCAGCTTCATGCTTAAGCTTGGAAACA 574
QY 300 GGTGTCCAAAATTTGCCAAGCATCTACCTGGAAGAGCTGATATAGATCAAGAACTATT 359
DB 575 GGTGTGCAAAATTTGCAAGCATCTTCGCGGAACAAGCAAGAGATTAAGAACTTCT 634
QY 360 GGAGACACAGATCCAGACACATCAA 387
DB 635 GGAGACAAAAGATTCCAGACACATGAA 662

RESULT 7

AAH87724
ID AAH87724 standard; cDNA; 626 BP.
XX
AC AAH87724;
XX
DT 25-SEP-2001 (first entry)
XX
DE Peppermint plant oil gland expressed cDNA 80.
XX
KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;
XX genetic mapping; antisense suppression; recombinant expression; ss.
OS Mentha x piperita.
XX
PN WO200153319-A1.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US02567.
XX
PR 20-JAN-2000; 2000US-0177264.
XX
PA (CROT/) CROTEAU R. B.
PA (LANGE/) LANGE B. M. R.
PA (WILD/) WILDUNG M. R.
XX
PI Croteau RB, Lange BM, Wildung MR;
XX
DR WPI; 2001-488706/53.
XX
PT New nucleic acid molecules corresponding to mRNA molecules expressed in
XX peppermint oil glands for enhancing expression of plant oil gland cell
XX proteins -
XX
PS Claim 1; Page 108; 251pp; English.
XX
CC The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
CC correspond to all or part of a mRNA molecule expressed in plant oil
CC gland cells, especially peppermint and plant oil glands that produce
CC terpenoid essential oils and resins. The nucleic acids are useful for
CC genetically mapping a plant genome for genes expressed in plant oil
CC gland cells and to suppress (for example by antisense suppression) or
CC enhance their expression (for example by genetically transforming a
CC plant cell with a replicable expression vector that expresses one or more
CC proteins naturally expressed in plant oil gland cells). The nucleic acids
CC are also useful for recombinant expression of plant oil gland proteins
CC required for terpenoid essential oil and/or resin production in bacterial
CC and/or yeast cells.
CC
SQ Sequence 626 BP; 187 A; 158 C; 163 G; 118 T; 0 other;
XX
Query Match 20.3%; Score 175; DB 22; Length 626;
Best Local Similarity 72.7%; Pred. No. 6.6e-26;
Matches 226; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 68 GAAGTGAAGAAAGGCGCTTGACATGGAAGAAAGCTTGATCTTGATGAACTATATGCA 127
DB 155 GAGGTGGAAGAAAGGCGCTTGACATGGAAGAAAGCTTGATCTTGATGAACTATATGCA 214
QY 128 AATCATGGGGAAGGTGTTGGAACCTTTGGCCAAAGCGTGTGTCGTAAGGTAACGA 187
DB 215 CATCAGGCGCAAGCCGAGTGAAGCTCTGCGCTCGCTCAGCAGGCTCAACGAAGTGA 274
QY 188 AAGAGTCCCGGCTAAGTGCTAAATTAACCTCCGTCGTGATGTTAGAAAGGGAATATT 247
DB 275 AAGAGTCCCGGCTAAGTGCTAAATTAACCTCCGTCGTGATGTTAGAAAGGGAATATT 334
QY 248 ACACCCGAGGAACAATTTGATATGAGAGCTCCAGCAAGTGGGGAACAGGTGTC 307
DB 335 ACTCTTGAAGAGCAGCTTTGATTCGACCTTCGATGAGGCGCAACAGGTGTC 394
QY 308 AAAATTGCCAAGCATCTACCTGGAAGACTGATATAGATCAAGAACTATTGGAGACA 367

The present sequence is *Arabidopsis thaliana* transcription factor, G1324 cDNA. The transcription factor is used for altering a plant's biochemical characteristics. The transcription factor may be used to alter the structure and developmental characteristics of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, roseaceous fruits and/or vegetable brassicas. Transcription factors are key controlling elements of biological pathways and altering expression levels of 1 or more transcription factors can change entire biological

AA WPT. 1999-
DB

RESULT 9	
ID	AAK25572
	AAK25572 standard; cDNA; 837 BP.
AC	AAK25572;
AD	02-AUG-1999 (first entry)
DE	Arabidopsis thaliana MYB2 (AtMYB2) cDNA.
XX	MYB2 gene; AtMYB2 gene; transcription factor; transgenic plant;
XX	environmental stress; stress tolerance; anaerobic stress;
XX	drought; flooding; salt; cold; crop protection; ss.
OS	Arabidopsis thaliana.
XX	
EH	Key
FT	Location/Qualifiers
FT	10..831
XX	/*tag= a
PN	W09916878-A1.
XX	
PD	08-APR-1999.
XX	
PF	25-SEP-1998; 98WO-AU00812.
XX	
PR	26-SEP-1997; 97AU-0009479.
XX	
PA	(CSTR) COMMONWEALTH SCI & IND RES ORG.
PI	Dennis ES, Dolferus RAM, Hoeren FU, Peacock WJ;
XX	WPI; 1999-263695/22.

DR P-PSDB: AAY05831.

XX Altering expression of a stress-related gene in a plant
XX Claim 22; Page 72-74; 119pp; English.

XX This is the sequence of Arabidopsis thaliana ecotype C24 cDNA
XX encoding the stress-related protein AtMYB2 (see AAY05831). The
XX cDNA was obtained by RT-PCR of anaerobically induced root RNA
XX derived from ecotype C24. The inventors have demonstrated that a
XX family of transcription factors, the MYB2 family of proteins, are
XX capable of inducing the expression of stress-related genes in plant
XX cells in response to stress situations. The MYB2 protein encoded
XX by the AtMYB2 gene is useful for the purposes of inducing or
XX repressing the expression of plant-expressible genes that are
XX involved in the plant's response to anaerobic stress, flooding
XX stress, cold stress, dehydration stress, drought stress, heat
XX stress or salinity. The stress-related gene, such as the MYB2
XX alcohol dehydrogenase Adh1 gene, is transactivated by the MYB2
XX cis-acting regulatory sequence, in particular a MB5-1 (Myb binding
XX site) motif and/or a GC-motif and/or a G-box-1 motif and/or a
XX G-box-2 motif, as found in the Adh1 promoter (see AAX25573-74). The
XX MYB2 protein is able to target several stress-induced enzymes,
XX rather than targeting just one, allowing the production of stress
XX tolerant transgenic plants.

SQ Sequence 837 BP; 246 A; 176 C; 212 G; 203 T; 0 other;

Query Match 19.6%; Score 169.4; DB 20; Length 837;
Best Local Similarity 69.5%; Pred. No. 8.8e-25;
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 60 AAGATCTGAGAGAGAAAGGCGCTTGACAAATGGAAGAAGCTTGATGTAAGT 119
DB 56 AAGATCTGATGACGGAAGAGTCCATGACCGAAGAAGATGCAATCTAGTCAACT 115
QY 120 ATATTGCAATGATGGGAGGTGTTGGAACCTTTGGCCAAGCTGCTGCTCAAC 179
DB 116 TCGCTCTATTTCATGCGATGCTGCTGGAACACATCGCTCTTCTCGGCTAAAGC 175
QY 180 GTAACGGAAGAGTGGCGGCTAAGTGCTAAATTAATCTCCGCTGATAGAGAAG 239
DB 176 GAAGTGGTAGAGTTGATGATTAAGTATGCTTAATTAATCTCCAGATGTAAGAAG 235
QY 240 GGAATATTACCCGAGGAACAATTGATTAATGAGCTCCACGCAAAAGTGGGAACA 299
DB 236 GCAACATCACTCTCGAAGAACATTAATGATCCCAAACTCTTTGGGCAATA 295
QY 300 GGTGCTCAAAATTTGCCAAGCATCTACCTGGAAGAGATGATGATCAAACTATT 359
DB 296 GTGTGTCGAAGATTGCCCATATCTACCGGAAGAACAGATTAATGAATTAAT 355
QY 360 GGAAGACAAAGATCCAGAACACATCAAGA 390
DB 356 GGAGAACTCGATCCAAAGCAAGCAACA 386

RESULT 10

ABK65252
ID ABK65252 standard; cDNA; 981 BP.

AC ABK65252;

DT 02-JUL-2002 (first entry)

DE Arabidopsis cDNA encoding a transcription factor #104.

XX Plant; ss; gene; transcription factor; transgenic;
XX agriculture; metabolic chemical; environmental stress; drought;
XX microbial disease resistance; herbicide resistance; seed yield;
XX fruit yield; growth rate; leaf senescence; flower senescence.

OS Arabidopsis thaliana.
XX WO200215675-A1.

XX 28-FEB-2002.

XX 22-AUG-2001; 2001WO-US26189.

XX 22-AUG-2000; 2000US-227439P.

XX 16-NOV-2000; 2000US-0713994.

XX 16-APR-2001; 2001US-0837944.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX (PIG/) PILGRIM M.

XX (CREE/) CREELMAN R.

XX (DUBE/) DUBBEL A J.

XX (HEAR/) HEARD J.

XX (JIAN/) JIANG C.

XX (KEDD/) KEDDIE J.

XX (ADAM/) ADAM L.

XX (RATC/) RATCLIFF O.

XX (REUB/) REUBER J L.

XX (RICH/) RIECHMANN J L.

XX (YU G/) YU G.

XX (PINE/) PINEDA O.

XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J,

XX Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;

XX WPI: 2002-292022/33.

XX P-PSDB: AAD93066.

XX An isolated or recombinant polynucleotide used to produce a transgenic

XX plant -

XX Claim 4; Page 443-445; 941pp; English.

XX The invention relates to 1 of 232 isolated or recombinant polynucleotides
XX encoding an Arabidopsis thaliana transcription factor, their variants,
XX complements, fragments, or related polynucleotide with 31% to 95%
XX sequence identity, where the plant possesses an altered trait as compared
XX to a wild-type or reference plant, or the plant exhibits an altered
XX phenotype as compared to a wild-type or reference plant, or the plant
XX exhibits ectopic expression or altered expression of one or more genes
XX associated with a plant trait as compared to a wild plant. Also included
XX are a transgenic plant comprising the polynucleotides, a computer
XX readable medium having stored sequence information, and identifying a
XX homologue sequence from a database comprising a plurality of known plant
XX sequences comprising inputting sequence information selected from one of
XX 464 fully defined sequences given in the specification. The isolated or
XX recombinant polynucleotide is used for producing a plant having a
XX modified trait, the method comprising selecting a polynucleotide that
XX encodes a polypeptide or an antisense nucleic acid, inserting the
XX polynucleotide or antisense nucleic acid into an expression vector,
XX introducing the vector into a plant or a cell of a plant to overexpress
XX the polypeptide or antisense nucleic acid, thereby producing a modified
XX plant, and selecting for a modified trait (e.g. increased
XX production of agriculturally useful proteins or metabolic chemicals,
XX pest tolerance, environmental stress response (e.g. drought), microbial
XX disease resistance, herbicide resistance, seed and fruit yield, growth
XX rate, leaf and flower senescence and many other traits listed in the
XX specification). The present sequence is one of the 232 polynucleotides
XX encoding an A. thaliana transcription factor.

SQ Sequence 981 BP; 319 A; 193 C; 235 G; 234 T; 0 other;

Query Match 19.6%; Score 169.4; DB 24; Length 981;
Best Local Similarity 69.5%; Pred. No. 9e-25;
Matches 230; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 60 AAGATCTGAGAGAGAAAGGCGCTTGACAAATGGAAGAAGCTTGATGTAAGT 119
DB 100 AAGAAATGACGTAGAGAGAGTCCATGACAGTTGAGGAAGATTAGAGCTCAAT 159

homectic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2

PA (YUGG/) YU G.

PA (PINE/) PINEDA O.
XX
PI Pilgrim M, Creelman R, Dubel J, Heard J, Jiang C, Keddle J;
PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
XX
DR WPI: 2002-292022/33.
DR P-PSDB: AAU93178.
XX
PT An isolated or recombinant polynucleotide used to produce a transgenic
PT plant -
XX
PS Claim 4; Page 881-883; 941pp: English.
XX
CC The invention relates to 1 of 232 isolated or recombinant polynucleotides
CC encoding an Arabidopsis thaliana transcription factor, their variants,
CC complements, fragments, or related polynucleotide with 31% to 95%
CC sequence identity, where the plant possesses an altered trait as compared
CC to a wild-type or reference plant, or the plant exhibits an altered
CC phenotype as compared to a wild-type or reference plant, or the plant
CC exhibits ectopic expression or altered expression of one or more genes
CC associated with a plant trait as compared to a wild plant. Also included
CC are a transgenic plant comprising the polynucleotides, a computer
CC readable medium having stored sequence information, and identifying a
CC homologue sequence from a database comprising a plurality of known plant
CC sequences comprising inputting sequence information selected from one of
CC 464 fully defined sequences given in the specification. The isolated or
CC recombinant polynucleotide is used for producing a plant having a
CC modified trait, the method comprising selecting a polynucleotide that
CC encodes a polypeptide or an antisense nucleic acid, inserting the
CC polynucleotide or antisense nucleic acid into an expression vector,
CC introducing the vector into a plant or a cell of a plant to overexpress
CC the polypeptide or antisense nucleic acid, thereby producing a modified
CC plant, and selecting for a modified trait (e.g. increased
CC production of agriculturally useful proteins or metabolic chemicals,
CC pest tolerance, environmental stress response (e.g. drought), microbial
CC disease resistance, herbicide resistance, seed and fruit yield, growth
CC rate, leaf and flower senescence and many other traits listed in the
CC specification). The present sequence is one of the 232 polynucleotides
CC encoding an A. thaliana transcription factor.
XX
SQ Sequence 972 BP; 318 A; 217 C; 216 G; 221 T; 0 other;
XX
Query Match 18.9%; Score 163.4; DB 24; Length 972;
Best Local Similarity 67.4%; Pred. No. 1.4e-23;
Matches 230; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 50 AAGAGCTCTCAAGATCTCTGAAGTGAAGAAAGCGCTTGACCAATGGAAGAACTTATC 109
DB 34 AACATGGAAGCGAGTGAAGCTAAAGAGAGTCCGTGAGCTGTGAAGAAAGATTTAAG 93
QY 110 TTGATGAACATATATGCAAAATCATGGGCAAGGTGTTGCAACTCTTGGCCAAAGCTGCT 169
DB 94 CTCATGAATTCATTGACTACTAATGAGAGAGGTGCTGCAACTCTCTTCGTTGGCCG 153
QY 170 GGTCCTCAACCTAAGCGAAGAGAGTGGCGGTGAAGTGGCTTAATTCCTCCGCTGAT 229
DB 154 GGGCTCCAAACCCACCGGTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 213
QY 230 GTTGAAGAGGGAATATTTACACCCGAGGAGCAACTTTTGTATGATGAGCTCCAGCAAG 289
DB 214 GTCCGCGGTGAAACATTACCTGTTGAAGAACACTCTTGATCTCGAATCTCATTTCCCGT 273
QY 290 TGGGGAAGACAGTGTCCAAATTTGCCAAGCATCTACCTGGAAGAGCTGATTAATGAGATC 349
DB 274 TGGGGAATATAGATGTGTAATAATGCACAATATTTACCGGGAAGAGCAAGCAAGAGATC 333
QY 350 AAGAATCTTTGGAGGACAAAGATCCAGAAGCATCAAGCA 390
DB 334 AAGAATCTAGTGAAGAGCGCGGTGCAAAAGCATCGCAACCA 374

ID AAC56152 standard; DNA; 389 BP.
XX
AC AAC56152;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #283.
XX
XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MIB; ss.
XX
OS Eucalyptus grandis.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
XX
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI: 2000-579369/54.
XX
XX New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX
PS Claim 1; Page 120; 747pp: English.
XX
XX The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and EREBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.
XX
SQ Sequence 389 BP; 112 A; 102 C; 97 G; 78 T; 0 other;
XX
Query Match 18.7%; Score 161.6; DB 21; Length 389;
Best Local Similarity 71.6%; Pred. No. 2.9e-23;
Matches 212; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 8 CTCATCACAACAAGATCATGATTAATAAACAACAGTGAAGCTCTCAAGATCCT 67
DB 29 CTCTCCAAAGTGAACATGACAAAGCAAGAGAGAGAGAGTGTATGCCAAGATGTC 88
QY 68 GAAGTGAGAAAGAGCGCTTGACCAATGGAAGAGAGCTTGATGATGATATATGCA 127
DB 89 GAAGTGAGAAAGAGCGCGTGAAGAGATGGAAGAGAGATCTCATCTCAACTACATAGG 148
QY 128 AATCATGAGGGAAGGTGTTGGAACCTCTTGGCCAAAGTGTGCTGCTCAACGTAACGCA 187
DB 149 AATCAGCGGCAAGAGCACTTGGAACCTCCTAAGCCAAAGCTGTGCTTAAGAGTACCGGG 208
QY 188 AAGAGTTGCGGCTCCGAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 247
DB 209 AAGAGTTGCGGCTCCGAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 268

1
2
3
4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Gencore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 07:57:25 ; Search time 1682 Seconds

(without alignments)
8309.575 Million cell updates/sec

Title: US-10-021-811-35

Perfect score: 863
Sequence: 1 gcaagagctctatcacac.....aaaaaaaaaaaaaaaa 863

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estrc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estlom:*
17: gb_gsa:*
18: em_gsa_hum:*
19: em_gsa_lnv:*
20: em_gsa_pln:*
21: em_gsa_vrt:*
22: em_gsa_fun:*
23: em_gsa_mam:*
24: em_gsa_mus:*
25: em_gsa_other:*
26: em_gsa_pro:*
27: em_gsa_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	65.1	562	14	BM732121 sa173b11.
2	506.4	58.7	586	12	BE820766 GM700012A
3	491.4	56.9	782	10	BE658316 BE527606
4	462.6	53.6	501	13	BM527606 sa163q06
5	444.4	51.5	450	10	BE058947 sn23b01.Y
6	442.8	51.3	523	9	AI930997 sb45h07.Y

7	425.4	49.3	431	10	AM459279 sh22h08.Y
8	424.4	49.2	552	13	BM527774 sa165g03.
9	401.2	46.5	568	13	BM527664 sa164q10.
10	400.6	46.4	544	13	BM527508 sa162f08.
11	369.8	42.9	409	14	BE804790 sa45f11.Y
12	369.6	42.8	560	14	BM732539 sa178e08.
13	368.6	42.7	635	14	BO146360 NF047E06F
14	363.6	42.1	476	10	BE057370 sm2c04.Y
15	347	40.2	542	13	BM528383 sa157f09.
16	330.8	38.3	425	10	AM423958 sh8e05.Y
17	307.8	35.7	682	14	BO146831 NF028G06F
18	306.2	35.5	688	13	BI272897 NF091A12F
19	304.8	35.3	669	12	BC457971 NF037A10P
20	304.6	35.3	395	12	BE805071 ss37a07.Y
21	294	34.1	470	10	AM156542 se28d11.Y
22	291.2	33.7	447	13	BI273011 NF097F02F
23	288.4	33.4	558	10	BE324639 NF024C04P
24	284	32.9	766	10	BE659054 GM700008A
25	279.4	32.4	715	10	BE658737 GM700007A
26	279	32.3	567	14	BM732162 sa173g09.
27	275	31.9	592	10	AM928296 EST307029
28	271.6	31.5	500	9	AI897784 EST267227
29	271.6	31.5	519	9	AI897681 EST267124
30	265.6	30.8	397	10	AM432364 sh73a04.Y
31	265.4	30.8	681	9	AI486576 EST244897
32	258.8	30.0	503	14	BO104458 gq1307.e
33	246	28.5	502	14	BO106505 f0056.e
34	240.2	27.8	463	9	AI938482 sb46a08.Y
35	223.8	25.9	516	12	BG652172 sad75b05.
36	219.2	25.4	467	12	BE807621 ss28c05.Y
37	217.8	25.2	370	12	BO642441 EST355917
38	217.8	25.2	371	9	AI487923 EST246245
39	204.8	23.7	542	9	AU238946 AU238946
40	193	22.4	492	13	BI262794 NF094E02E
41	192.4	22.3	480	10	BE023108 sm9h01.Y
42	190.2	22.0	343	12	BE804882 ss34c01.Y
43	189.4	21.9	256	12	BE805198 ss39d09.Y
44	184.8	21.4	510	12	BF325282 su20e03.Y
45	183.6	21.3	295	13	BM527578 sa163e01.

ALIGNMENTS

RESULT 1
BM732121 562 bp mRNA linear EST 01-MAR-2002
LOCUS sa173b11.Y1 Gm-cl061 glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl061-4678 5' similar to YR:Q39028 Q39028 ATMYB2. ; mRNA
sequence.

ACCESSION BM732121
VERSION BM732121.1 GI:19053454
KEYWORDS EST.

SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 562)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V., Rhana
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact: cou@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Glibco
 High quality sequence stop: 421.
 Location/Qualifiers

FEATURES

source

1. 562
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl061-4678"
 /clone_1lb="Gm-cl061"
 /tissue_type="mature flowers of field grown plants"
 /lab_host="DH10B"
 /note="Vector: Bluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from mature flowers of field grown plants for the cultivar
 Raiden. Complementary DNA was synthesized from mRNA using
 a primer consisting of a poly(dT) sequence with a XhoI
 restriction site. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the Bluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."

BASE COUNT

187 a 115 c 128 g 132 t

ORIGIN

Query Match 65.1%; Score 562; DB 14; Length 562;
 Best Local Similarity 100.0%; Pred. NO. 2.8e-74;
 Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

10 CTATCACACACACAGTCATGATGATTAACAAACAGTGAAGAGCTCAAGATCTCTGA 69
 1 CTATCACACACACAGTCATGATGATTAACAAACAGTGAAGAGCTCTCAAGATCTCTGA 60
 70 AGTGGAGAGAGGCGCTTGACATGAGAGAGAGCTGATGATGATGATGATGATGATGATG 129
 61 AGTGGAGAGAGGCGCTTGACATGAGAGAGAGCTGATGATGATGATGATGATGATGATG 120
 130 TCATGGGAGAGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 189
 121 TCATGGGAGAGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 190 GAGTGGCGGCTGAAGTGAATTAATCTCCGCTGATGATGATGATGATGATGATGATGATG 249
 181 GAGTGGCGGCTGAAGTGAATTAATCTCCGCTGATGATGATGATGATGATGATGATGATG 240
 250 ACCGAGAGACACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 309
 241 ACCGAGAGACACTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 310 AATGCGAGAGATCTACCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 369
 301 AATGCGAGAGATCTACCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 360
 370 GATCCAGAGACATCAAGCAAGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 429
 361 GATCCAGAGACATCAAGCAAGCTGAGAGATCTGAGAGATCTGAGAGATCTGAGAGAT 420
 430 GATAAATATCAACCAAGCTAGACATGATGATGATGATGATGATGATGATGATGATGAT 489
 421 GATAAATATCAACCAAGCTAGACATGATGATGATGATGATGATGATGATGATGATGAT 480
 490 GATATCTCCACCTGTATCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 549
 481 GATATCTCCACCTGTATCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
 550 TAATCTGATCAATCACTGTGT 571
 |||||||

Db 541 TAATCTGATCAATCACTGTGT 562

RESULT 2

BE820766/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1. 586

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="Gm-r1070-4966"

/clone_1lb="Gm-r1070"

/note="The library Gm-r1070 is a sequence-driven, rerecked

set of 9,216 clones selected from cDNA libraries from

various tissues and stages of development of soybean that

represent 2,639 sequences from immature cotyledons, 1,770

from immature seed coats, 3,938 from flowers, and 869

from young pods. The 5' ESTs of the source clones from

the different libraries was used to select singletons, or

a representative of each contig, which were rerecked to

form library Gm-r1070. The cDNA clones of the rerecked

Gm-r1070 library were then sequenced at the 3' end. The

contig analysis to select unique genes was performed by

the laboratory of Ernest Ritzel, Center for Computational

Genomics and Bioinformatics, University of Minnesota,

http://www.cbc.umn.edu/ResearchProjects/soybean/index.html

Rerecked was performed by Genome Systems, St. Louis,

http://www.genomesystems.com, and 3' sequencing by the

Keck Center for Comparative and Functional Genomics,

University of Illinois,

http://www.life.uiuc.edu/biotech/keck.html. Note: The

corresponding 5' EST from each clone in the Gm-r1070

library is listed in the 'OTHER EST' field. The detailed

information on the source library for each clone can also

be obtained by referring to the Genome Systems clone ID of

'OTHER EST'."

OTHER EST"

173 a 110 c 100 g 194 t 9 others

58.7%; Score 506.4; DB 12; Length 586;

Query Match

Db	Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	302	ACAAAGAGATTAACTATGTGAGCATGAGGAGTATCTGGTCAATGATAGTACTGTAACGGG	243							
Db	644	GATTAAATAT-----TATCAGATATAAACCAATAGTTCG---AAGTTCAT--AAGCGTG	693							
Db	242	GATTAAATATGTGATATATCAAGATTAACCTAATATCTGTATTAAGTTCATTAACCACTG	183							
Db	694	GAATGCTCTGCGCTTAAACATATATATTAGGTTGTTATATTAAGTACTGG--ATG	748							
Db	182	GAATGCTCTGCGCTTAAACATATATATTAGGTTGTTATATTAAGTACTGGATATG	123							
Db	749	TTTGCTTTGGCTACCATTTATAGC 773								
Db	122	TTTGCTTTGGCTACCATTTATAGC 98								
RESULT 4	BMS27606	501 bp	mrna	linear	EST 19-FEB-2002					
LOCUS	BMS27606	501 bp	mrna	linear	EST 19-FEB-2002					
DEFINITION	BMS27606.y1 Gm-cl061 glycine max cDNA clone SOYBEAN	501 bp	mrna	linear	EST 19-FEB-2002					
ACCESSION	BMS27606	501 bp	mrna	linear	EST 19-FEB-2002					
VERSION	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
KEYWORDS	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
SOURCE	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
ORGANISM	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
REFERENCE	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
AUTHORS	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					
COMMENT	BMS27606.1	501 bp	mrna	linear	EST 19-FEB-2002					

	ORIGIN	BASE COUNT	170 a	104 c	114 g	113 t	
Query Match	Best Local Similarity	53.6%;	Score 462.6;	DB 13;	Length 501;		
Matches 477;	Conservative	95.2%;	Pred. No. 1.5e-59;	Mismatches 24;	Indels 0;	Gaps 0;	
OY	20	CACAAATCATGATATAAAAAACAACAGCTGTAAAGACGTCCTCAAGATCCGAAAGTGAAAAA	79				
Db	1	CACAAAACAATGCATTAACAAAAAACACTGTGGCAACACGTCCTCAAGATCCGAAAGTGAAAAA	60				
OY	80	GGGCCCTTGGACAAATGAAGAAGACTTGTATTCTGTATGACTATATATGCAAAATCATGGGAAA	139				
Db	61	GGACCTTGGACATGGAAGAAGACTTTGATCCTTGATCACTATATTGCAAAATCATGGGAAA	120				
OY	140	GGTGTGTGGAAGCTCTTTGGGCCAAAGCTGTGTCTCAAAACGTAACGGAAGAAGTTGCCGG	199				
Db	121	GGTGTGTGGAATTCCTTTGGCCCAAAGCTGTGTCTCAAAACGTAACGGAAGAAGTTGCCGG	180				
OY	200	CTAAGGTGGCTAAATTAACCTCCGCTCGTAGTTAGAAAGAGGAATATTACACCCGAGAAA	259				
Db	181	CTAAGGTGGCTAAATTAACCTCCGCTCGTAGTTAGAAAGAGGAATATTACACCCGAGAAA	240				
OY	260	CAACTTTTGATATATGAGAGCTCCAGCAAAAGTGGGGAAACAGTGTGTCGCAAAATTTGCCAAG	319				
Db	241	CAACTTTTGATATATGAGAGCTTCACCAGAAAGTGGGGAAACAGTGTGTCGCAAAATTTGCCAAG	300				
OY	320	CATCTTACTGTGAAGACATGATATATGATCAAGAAGTATTGACGACAAAGATCCAGAAAG	379				
Db	301	CATCTTACTGTGAAGACATGATATATGATCAAGAAGTATTGACGACAAAGATCCAGAAAG	360				
OY	380	CACATCAAGCAAGCTGAGAACTTTCAGCAACAGAGTAGTAATTAATCTCGATATAATGAT	439				
Db	361	CACATCAAGCAAGCTGAGAACTTTCAGCAACAAATTAATTAATCTCGATATAATGAT	420				
OY	440	CACCAAGCTGAGACATAGCCATGTTTCCACATGAGCGTGCAGACCCCATTGAGATGATATCTCCA	499				
Db	421	CACCAAGCTGAGACATAGCCATGTTTCTCCATGAGCGTGCAGAACCCATTGATATCTCCA	480				
OY	500	CCCTGTTATCAAGGAATGTTA 520					
Db	481	CCCTTTTATCAAGGAATGTTA 501					
RESULT 5	BE058947	450 bp mRNA linear EST 03-DEC-2001					
LOCUS	sn23b01.y1	Gm-cl01016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID					
DEFINITION	Gm-cl01016-12290.5,	similar to TR-09ZTD5 Q9ZTD5 PUTATIVE					
TRANSCRIPTION FACTOR ;	mRNA sequence.						
ACCESSION	BE058947						
VERSION	BE058947.1	GI:8403313					
KEYWORDS	EST.						
SOURCE	Glycine max						
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;						
	Glycine.						
REFERENCE	1 (bases 1 to 450)						
AUTHORS	Shoenaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corvelli,V., Khanna,A., Bolla,B., Meira,M., Hillier,L., Kudaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,B., Thelning,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterson,R. and Willson,R.						
TITLE	Public Soybean EST Project						
JOURNAL	Unpublished (1999)						
COMMENT	Contact: Shoenaker R/public soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800						

Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 411.

FEATURES

source

Location/Qualifiers
 1..450
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-12290"
 /clone_1ib="Gm-c1016"
 /tissue_type="Immature flowers of field grown plants"
 /lab_host="X110-Gold"
 /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from immature flowers of field grown plants. The cDNA
 library was prepared using the Stratagene pBluescript II
 XR library construction kit. Complementary DNA was
 synthesized from mRNA using a primer consisting of a poly
 (dT) sequence with a XhoI restriction site. EcoRI adapters
 were ligated to the blunt-ended cDNA fragments followed by
 XhoI digestion. The cDNA fragments were directionally
 cloned into the EcoRI-XhoI restriction site of the
 pBluescript vector. The ligated cDNA fragments were
 transformed into X110-Gold host cells. This library was
 constructed by Dr. Randy Shoemaker and Dr. John
 Erpelting."

BASE COUNT 151 a 84 c 90 g 125 t
 ORIGIN

Query Match

Best Local Similarity 98.9%; Score 444.4; DB 10; Length 450;
 Matches 445; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY 304 GTCCAAAATGCCAGATCTACCTGGAGGAGTGTGAATGAGATCAAGACATTGGAG 363
DB 1 GTCCAAAATGCCAGATCTACCTGGAGGAGTGTGAATGAGATCAAGACATTGGAG 60
OY 364 GACAAAGATCCAGACACATCAAGCAAGCTGGAATCTTCAGCAACAGATGTAATA 423
DB 61 GACAAAGATCCAGACACATCAAGCAAGCTGGAATCTTCAGCAACAGATGTAATA 120
OY 424 TTCTGAGATTAATGATCACCAGCTAGCAGTATGTTCCACCATGGCTGAGCCAT 483
DB 121 TTCTGAGATTAATGATCACCAGCTAGCAGTATGTTCCACCATGGCTGAGCCAT 180
OY 484 GGAGATGATTTCCACCGCTGATCAAGGAATGTAGAGCCATTTTCAACTGATTTCC 543
DB 181 GGAGATGATTTCCACCGCTGATCAAGGAATGTAGAGCCATTTTCAACTGATTTCC 240
OY 544 TACAATTAATCTGATCAATCCAGTTGTTGTACCAATGACAAACACATTAATTTG 603
DB 241 TACAATTAATCTGATCAATCCAGTTGTTGTACCAATGACAAACACATTAATTTG 300
OY 604 GACCAATGAGGATAGCTGTGCAATGCAATTAAGCAAGGATTAATATATCAAGATA 663
DB 301 GACCAATGAGGATAGCTGTGCAATGCAATTAAGCAAGGATTAATATATCAAGATA 360
OY 664 AAACCTAAGTTTGAAGTTCCATAGCGTGAATGCTTGGATTAACATATTTTGG 723
DB 361 AAACCTAAGTTTGAAGTTCCATAGCGTGAATGCTTGGATTAACATATTTTGG 420
OY 724 GTTGTATTATTAAGTAGTTGATGTTGG 753
DB 421 GTTGTATTATTAAGTAGTTGATGTTGG 450

```

RESULT 6
 LOCUS AT930997 523 bp mRNA linear EST 30-NOV-2001
 DEFINITION sb45h07.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

ACCESSION

AT930997.1 GI:5666961

VERSION

EST.

KEYWORDS

soybean.

SOURCE

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 523)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khana

A., Bolla,B., Merritt,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

Y., Pearson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk

R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann

R., Mesterion,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 422.

Location/Qualifiers

1..523

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-278"

/clone_1ib="Gm-c1015"

/tissue_type="Mature flowers, field grown plants"

/lab_host="X110-Gold"

/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:

XhoI; This cDNA library was constructed from mRNA isolated

from mature flowers of field grown plants. The cDNA

library was prepared using the Stratagene pBluescript II

XR cDNA library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pBluescript vector. The ligated cDNA fragments were

transformed into X110-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Erpelting."

BASE COUNT 172 a 124 c 111 g 115 t 1 others

ORIGIN

Query Match 51.3%; Score 442.8; DB 9; Length 523;
 Best Local Similarity 94.3%; Pred. No. 1.3e-56;
 Matches 459; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

OY 8 CTCATACACACACAGTCAATGGATTAATAAACAACAGTGAAGCTCTCAAGATCCT 67
DB 37 CTCATACACACACACAGTCAATGGATTAATAAACAACAGTGAAGCTCTCAAGATCCT 96
OY 68 GAAGTGAAGAAAGGCGCTTGACATGGAAGAAGCTTGATGATGATGATGATGATGATG 127
DB 97 GAAGTGAAGAAAGGCGCTTGACATGGAAGAAGCTTGATGATGATGATGATGATGATG 156
OY 128 AATCATGGGAAGTGTGTTGGAAGCTTTGGCCAAAGCTGTGCTGCTCAACGTAAGGA 187
DB 157 AATCATGGGAAGTGTGTTGGAAGCTTTGGCCAAAGCTGTGCTGCTCAACGTAAGGA 216

```



```

DEFINITION      ss45f11.y1 Gm-c1061 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
                  Gm-c1061-1702 5' similar to TR:039028 Q39028 ATMB2. ;, mRNA
ACCESSION       BE804790
VERSION         BE804790.1 GI:10235902
KEYWORDS
SOURCE          soybean.
ORGANISM        Glycine max
                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                 Glycine
REFERENCE       1 (bases 1 to 409)
AUTHORS        Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
                 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
                 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
                 R., Waterston,R. and Wilson,R.
JOURNAL        Public Soybean EST Project
COMMENT        Unpublished (1999)
                 Contact: Shoemaker R/Public Soybean EST Project
                 Washington University School of Medicine
                 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                 Tel: 314 286 1800
                 Fax: 314 286 1810
                 Email: est@watson.wustl.edu
                 This clone is available through: Resgen, Invitrogen Corp. 2130
                 South Memorial Parkway Huntville, AL 35801 For further information
                 call: (800)-533-4363 or contact via email: ccu@resgen.com
                 Insert Length: 1068 Std Error: 0.00.
FEATURES
  source
    1..409
    /organism="Glycine max"
    /db_xref="taxon:3847"
    /clone="GENOME SYSTEMS CLONE ID: Gm-c1061-1702"
    /clone_1lb="Gm-c1061"
    /tissue_type="mature flowers of field grown plants"
    /lab_host="DH10B"
    /note="Vector: Bluescript II SK+; Site:1: EcoRI; Site:2:
    XhoI; The cDNA library was constructed from mRNA isolated
    from mature flowers of field grown plants for the cultivar
    Raiden. Complementary DNA was synthesized from mRNA using
    a primer consisting of a poly(dT) sequence with a XhoI
    restriction site. EcoRI adapters were ligated to the
    blunt-ended cDNA fragments followed by XhoI digestion. The
    cDNA fragments were directionally cloned into the
    EcoRI-XhoI restriction site of the Bluescript vector. The
    ligated cDNA fragments were transformed into DH10B host
    cells (GibcoBRL). This library was constructed in the
    laboratory of Dr. Randy Shoemaker."
BASE COUNT      136 a      87 c      97 g      89 t
ORIGIN
Query Match     42.9%; Score 369.8; DB 12; Length 409;
Best Local Similarity 95.7%; Pred. No. 9.4e-6;
Matches 380; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

Db 193 AACAGTCCGCGGCTAAGGTGGCTAAACTACCTCCCTCGATGATTGAAGACGAATATT 252
Qy 248 ACACCCGAGAACAACTTTTGAATTATGAGCTCCAGCAAGTGGGAAACAGGTGTC 307
Db 253 ACACCCGAGAACAACTTTTGAATTATGAGCTCCAGCAAGTGGGAAACAGGTGTC 312
Qy 308 AAAATTGCCAAGCATCTACCTGGAAGACTGATTAATGATCAAGACTATTGGACACA 367
Db 313 AAAATTGCCAAGCATCTACCTGGAAGACTGATTAATGATCAAGACTATTGGAGACC 372
Qy 368 AGGATCCAGACACATCAAGCAAGCTGGAACCTTC 404
Db 373 AGGATCCAGACACATCAAGCAAGCTGGAACCTTC 409
RESULT 12
LOCUS          BM732539
DEFINITION     BM732539 560 bp mRNA linear EST 01-MAR-2002
                 sal78e08.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
                 Gm-c1061-5223 5' similar to TR:Q39028 Q39028 ATMB2. ;, mRNA
                 sequence.
ACCESSION      BM732539
VERSION        BM732539
KEYWORDS
SOURCE         soybean.
ORGANISM       Glycine max
                 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                 Glycine
REFERENCE       1 (bases 1 to 560)
AUTHORS        Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
                 A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
                 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
                 R., Waterston,R. and Wilson,R.
JOURNAL        Public Soybean EST Project
COMMENT        Unpublished (1999)
                 Contact: Shoemaker R/Public Soybean EST Project
                 Washington University School of Medicine
                 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                 Tel: 314 286 1800
                 Fax: 314 286 1810
                 Email: est@watson.wustl.edu
                 This clone is available through: Resgen, Invitrogen Corp. 2130
                 South Memorial Parkway Huntville, AL 35801 For further information
                 call: (800)-533-4363 or contact: ccu@resgen.com web site:
                 www.resgen.com
                 Seq primer: -40RP from Gibco
                 High quality sequence stop: 430.
FEATURES
  source
    1..560
    /organism="Glycine max"
    /db_xref="taxon:3847"
    /clone="SOYBEAN CLONE ID: Gm-c1061-5223"
    /clone_1lb="Gm-c1061"
    /tissue_type="mature flowers of field grown plants"
    /lab_host="DH10B"
    /note="Vector: Bluescript II SK+; Site:1: EcoRI; Site:2:
    XhoI; The cDNA library was constructed from mRNA isolated
    from mature flowers of field grown plants for the cultivar
    Raiden. Complementary DNA was synthesized from mRNA using
    a primer consisting of a poly(dT) sequence with a XhoI
    restriction site. EcoRI adapters were ligated to the
    blunt-ended cDNA fragments followed by XhoI digestion. The
    cDNA fragments were directionally cloned into the
    EcoRI-XhoI restriction site of the Bluescript vector. The
    ligated cDNA fragments were transformed into DH10B host
    cells (GibcoBRL). This library was constructed in the
    laboratory of Dr. Randy Shoemaker."
BASE COUNT      191 a      128 c      111 g      130 t

```

ORIGIN

Query Match	42.8%;	Score 369.6;	DB 14;	Length 560;	
Best Local Similarity	80.0%;	Pred. No. 8.3e-46;			
Matches 448;	Conservative 0;	Mismatches 109;	Indels 3;	Gaps 1;	

FEATURES

Email:	gdmay@noble.org
Insert Length:	635
Std Error:	0.00
Plate:	047
row:	E
column:	06
Seq primer:	TCACACAGAAACACGCTGTGC.
Location/Qualifiers	

FEATURES
SOURCE

Y	54	CGCTCCAAAGTTCCTCAAGTACGAAAAAGGCGCTTGACATATGAGAAAGACTTCATCTTGA	113
Y	54	CGCTCCAAAGTTCCTCAAGTACGAAAAAGGCGCTTGACATATGAGAAAGACTTCATCTTGA	113
Db	4	CATCTCATGATCTCTGAAAGTGAAGAAAGGAGCCATGACCATGACATGAAAGACTTCATCTTGA	63
Y	114	TGAACATATATGGAATCATGGGAAAGGTGTTTGGAACTCTTTGGCCAAAGCTGTGTC	173
Y	114	TGAACATATATGGAATCATGGGAAAGGTGTTTGGAACTCTTTGGCCAAAGCTGTGTC	173
Db	64	TAAACATATATGGAATCAAGGTGAAGGTGTTTGGAACTCTTGAAGCCAAAGCTTGTGTC	123
Y	174	TCAAAGCTAACCGAAAGAGTTGCCGCTAAGGTGGCTAAATTACCTCGTCTGATGTTA	233
Y	174	TCAAAGCTAACCGAAAGAGTTGCCGCTAAGGTGGCTAAATTACCTCGTCTGATGTTA	233
Db	124	TCAAAGCAACGGGAAAGAGTTGTCGACTCCGTTGGCTAACTACCTTCGTCGATGTTA	183
Y	234	GAAAGAGGAATATTTACACCGAGAAACAACATTTGATATAGACCTCCAGCAAGTCGG	293
Y	234	GAAAGAGGAATATTTACACCGAGAAACAACATTTGATATAGACCTCCAGCAAGTCGG	293
Db	184	GAAAGAGGAATTTTACACCGAGAAACAACATTTGATATAGACCTTCATGCAAAAGTCGG	243
Y	294	GAAACAGGTGTGCCAAAATTTGCCAAGCATCTTACCTGGAAGACTGATATAGATCAAGA	353
Y	294	GAAACAGGTGTGCCAAAATTTGCCAAGCATCTTACCTGGAAGACTGATATAGATCAAGA	353
Db	244	GCAATAGGTGTGCCAAAATTTGCCAAGCATCTTCCAGGAAGCACTGAACATGAAATTAAGA	303
Y	354	ACATATTGAGAGCAAGATATCCAGAAACATCAACAGCACTGAGAACCTTCACAGACAGA	413
Y	354	ACATATTGAGAGCAAGATATCCAGAAACATCAACAGCACTGAGAACCTTCACAGACAGA	413
Db	304	ACTTTGGAGAACAAAGATATCCAAAAGCAATTAAGCAAGCTGAGAACTTCACAAACATAG	363
Y	414	GTAGTATATATTTGTGAGATAAATGATCACAAGCTAGACATAGCAGCATGTTTCCACCATG	473
Y	414	GTAGTATATATTTGTGAGATAAATGATCACAAGCTAGACATAGCAGCATGTTTCCACCATG	473
Db	364	GTATATTGAGAAATTAATGATCATCAAGCAAGCATAGTACGAAAGGTGCCACCATG	423
Y	474	CTGAGCCCATGAGATGATATCTCCACCCTGTATTAACAAGAAATGTGAAGCCATTTTCA	533
Y	474	CTGAGCCCATGAGATGATATCTCCACCCTGTATTAACAAGAAATGTGAAGCCATTTTCA	533
Db	424	CACATCCAATGAGATGATTTCTCTCCACCCTCATACCAAGAAATTTTGAAGCCATTTCA	483
Y	534	CTCAGTTCCTCAATTAATTAATCTGATCAATCCAGTTGTTGTAACCAATGACAAACACAA	593
Y	534	CTCAGTTCCTCAATTAATTAATCTGATCAATCCAGTTGTTGTAACCAATGACAAACACAA	593
Db	484	CTCAATTTCCCTACAAATCA---CTGATCAATCAAGTTGTTGTAACCAACCAACGACAGCA	540
Y	594	TTTAACATTTGGAGCATGGAG	613
Y	594	TTTAACATTTGGAGCATGGAG	613
Db	541	ACAACATTTGGAGCATGGAG	560
Db	541	ACAACATTTGGAGCATGGAG	560
RESULT	13		
LOCUS	BO146360	635 bp	linear
DEFINITION	NP047E06EFL1050	Developing flower	Medicago truncatula cDNA clone
ACCESSION	BO146360	5', mRNA sequence.	
VERSION	BO146360.1	GI:20283419	
KEYWORDS	EST.		
SOURCE	barrel medic.		
ORGANISM	Medicago truncatula		
REFERENCE	1 (bases 1 to 635)		
AUTHORS	Torres-Jerez, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,		
TITLE	Flowers, H.R., Inman, J.T., Weller, J.W. and May, G.D.		
JOURNAL	Expressed Sequence Tags from the Samuel Roberts Noble Foundation		
COMMENT	Medicago truncatula flower library		
	Unpublished (2001)		
	Contact: May GD		
	Plant Biology Division		
	The Samuel Roberts Noble Foundation		
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA		
	Tel: 580 221 7391		
	Fax: 580 221 7380		

RESULT 13

LOCUS	BQ146360	635 bp	MRNA	linear	EST 24-APR-2002
DEFINITION	NF047E06FLJ1050	Developing flower	Medicago truncatula	cDNA clone	
ACCESSION	NF047E06FLJ 5 ,	mRNA sequence.			
VERSION	BQ146360				
KEYWORDS	BQ146360.1	GI:20283419			
SOURCE	EST.				
ORGANISM	barrel medic.				
	Medicago truncatula				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;				
	Medicago.				
REFERENCE	1 (bases 1 to 635)				
AUTHORS	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,				
TITLE	Flores, H.R., Innem, J.T., Weller, J.W. and May, G.D.				
JOURNAL	Expressed Sequence Tags from the Samuel Roberts Noble Foundation				
COMMENT	Medicago truncatula Flower Library				
	Unpublished (2001)				
	Contact: May GD				
	Plant Biology Division				
	The Samuel Roberts Noble Foundation				
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA				
	Tel: 580 221 7391				
	Fax: 580 221 7380				

Query Match	Best Local Similarity	Matches 486;	Conservative	Score 368.6;	DB 14;	Length 635;
Qy 8	CTCTATCACACACACAGTCAATGGATATATAAAAAACACAGTGAAGAGTGTCAAGATCCT	42.7%;	78.0%;	0;	Mismatches 124;	Indels 13; Gaps 3.
Db 13	CTCTCATATTCACATATATATATAAAAAATATGGATATATAAAAAACCTCAACATCATCTCAAGATCCT					
Qy 68	GAAGTGAAGAAAGGCGCTTGACCAATGGAAGAGAGACTGATCTGATGACACTATATATGCA					
Db 73	GAAGTGAAGAAAGGCGCATGAGCACTGAGACATGGAAGAGAGCTGATTTATATCATATATATATGCA					
Qy 128	AATCATGGGGAAGTGTGGGAACCTTTGGCCAAAGTGTGTGTCTCAAAAGCTAACGGA					
Db 133	AATCATGGGGAAGTGTGGGAATCTCTTGAAGCTTGAAGCAAGCTGTGTGTCTTAAACGTACGGA					
Qy 188	AAGAGTTGCCGGGCTAAGGTGGCTAATTAATTAACCTCCGCTCTGTATGTTAGAAAGGGAATAT					
Db 193	AAAAGTTGCCAGGCTGTGATGTGTTAAACTTAACCTCTCTCCAGATGTTTAAAGAGGAATAT					
Qy 248	ACACCCGAGAAACAATTTGATTTATGAGCGTCCACGCAAGTGGGGAAGAAAGAGTGTCC					
Db 253	ACACCTGAGGAACAACCTTTGATGATCATGACATCTCATGCAAAAGTGGGGAATAGTGTGTC					
Qy 308	AAAATTGCCAGCATCTACCTGGAAGAGCTGATTAATGATCAAGCAAGAACTTATGGAGACA					
Db 313	AAAATTGCCAAGCATCTTCCAGAGAAAGCTGCAATGAGATTAAGAAATTTTGGAGACT					
Qy 368	AGATTCAGAGACATCAAGCAAGCTGGAATCTTCAGAACAGAGT-----AGTAAT					
Db 373	AGATTCAGAGACATCAAGCAAGCTGGAATCTTCAGAACAGAGT-----AGTAAT					
Qy 422	AATTCGAGATTAATGATCACAAGCTAGCACTAG-----CGATGTTTCCACATAGGCTGA					
Db 433	AGTTTGAAGATTAATGATCACAATCAATCAAGTATAGCCCAATTTTCCAAACATTTAGGA					
Qy 478	GCCCATGAGATTAATTCACACCTGTTTACAGAGAAATGTTAGAGCAATTTTCACTCA					
Db 493	GCCCATGAGATTAATTCACACCTGTTTACAGAGAAATGTTAGAGCAATTTTCACTCA					
Qy 538	GTTCCCTACAAAT---TAATCTGATCAATCCAGTGTGTACCAATGACAAACAACAAT					
Db 553	ATTTCACACATATGATCATCATCCCTCAAGTGTGTACCAATGACAAACAACAAT					
Qy 595	TAACATATTGAGACATGAGAGATA					

/tissue_type="mature flowers of field grown plants"
/lab_host="DH10B"
/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants for the cultivar
Raiden. Complementary DNA was synthesized from mRNA using
a primer consisting of a poly(dT) sequence with a XhoI
restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the Bluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."

BASE COUNT 191 a 121 c 107 g 123 t
ORIGIN

Query Match 40.2%; Score 347; DB 13; Length 542;

Best Local Similarity 78.3%; Pred. No. 1.9e-42;

Matches 416; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 28 AATGATATAAAACACAGCTGTAAGAGCTCTCAAGATCTGAAGTGAGAAAAGGCTTG 87
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 5 AATGACAAAAAACACAGCTGTAAGAGCTCTCAAGATCTGAAGTGAGAAAAGGCTTG 64
QY 88 GACATGGAAGAGACTGATCTGATGAATATATGCAAAATCATGGGAGGTTTG 147
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 65 GACCATGGAAGAGACTGATCTGATGAATATATGCAAAATCATGGGAGGTTTG 124
QY 148 GAATCTTTGGCCAAAGCTGCTGCTCAAAAGCTTAACGAAAGAGTTGCCGCTAAGGTG 207
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 125 GAATCTCTTGGCCAAAGCTGCTGCTCAAAAGCTTAACGAAAGAGTTGCCGCTAAGGTG 184
QY 208 GCTAAATTAACCTCCGCTCTGATGTTAGAGAGGGAATATTAACCCGAGGAACACTTTT 267
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 185 GCTAACTACCTGCTGCTGATGTTAGAGAGGGAATATTAACCCGAGGAACACTTTT 244
QY 268 GATTATGAGCTCCACGCAAGTGGGGAACAGTGGTCCAAAATTTGCCAAGCATCTACC 327
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 245 GATCATAGAACTTCATGCAAAAGTGGGCAATAGGTGGTCCAAAATTTGCCAAGCATCTTC 304
QY 328 TGAAGGACTGATTAATGATCAAGAACTATTGGAGAGCAAGATCCAGAACATCA 387
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 305 AGAAGAACTGACAAATGAGATTAGAACTTCTGGAGAAAGATCCAAAAGCATTTAA 364
QY 388 GCAAGCTGAGAACTTTGAGCAACAGATAGTAATTTCTGAGATAATGATCACCAAGC 447
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 365 GCAAGCTGAGAACTTCACAAACATGGAATTCAGAGAAATGATGATCAAGCAAGCAC 424
QY 448 TAGCAGTAGCCATGTTTCCACCATGCTGAGCCATGAGATGATTTCTCCACCCTGTTA 507
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 425 TAGTACTAGCAAAAGTGTCCACATGAGATCCAAATGAGACTTTCTCTCACCCCTCATA 484
QY 508 TCAAGGAATGTAGAGCATTTTCAACTCACTGCCCTACAAATTAATCTTGA 558
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 485 CCAAGGAACCTTTGAGCCATTCACCTCAATTCCTACATCATGATCA 535

Search completed: February 18, 2003, 09:19:17
Job time : 1690 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 18, 2003, 10:32:35 ; Search time 47 Seconds

(without alignments)
1344.158 million cell updates/sec

Title: US-10-021-811-36

Perfect score: 1138
Sequence: 1 MDRKQCKTSQDEVRKGPV.....NNINYSMEDSMWOLLNED 206

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+.p2n.model -DEV-x1h
-O=/cgn2_1/uspro.spool/US10021811/rnat.12022003.090134.9582/app_query.fasta_1.391
-DB-Issued_Patents_NA -OPMT-fastap -SUFFIX-p2n.rni -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human4.0.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTPMT=ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021811.0CGN_1.1_17-@rnat.12022003.090134.9582 -NCPV=6 -ICPV=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NDC_SCORES=0 -MAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
5: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
6: /cgn2_6/prodata/1/lna/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	396	34.8	2220 4	US-08-997-251-1
2	386	33.9	2352 4	US-08-997-251-1
3	382.5	33.6	1344 2	US-08-722-6268-1
4	305.5	26.8	1035 5	US-08-306-691B-45
5	305.5	26.8	3225 1	US-08-306-691B-45
6	305.5	26.8	3225 5	US-08-306-691B-45
7	305.5	26.8	3230 5	US-08-306-691B-45
8	296	26.0	1897 4	US-09-167-3222-10
9	293	25.7	4880 4	US-09-402-9225-5
10	291.5	25.6	3602 4	US-09-402-9225-5
11	291.5	25.6	6775 4	US-09-402-9225-5
12	271.5	23.9	2638 1	US-08-306-691B-46

13	248	21.8	4059 2	US-08-485-139-1	Sequence 1, Appl1
14	248	21.8	4059 3	US-08-750-357-1	Sequence 1, Appl1
15	248	21.8	4824 2	US-08-485-139-5	Sequence 5, Appl1
16	248	21.8	4824 3	US-08-750-357-5	Sequence 5, Appl1
17	185.5	16.3	469 3	US-09-008-979A-2	Sequence 2, Appl1
18	185.5	16.3	469 4	US-09-460-618-2	Sequence 2, Appl1
19	185.5	16.3	469 4	US-09-460-618-2	Sequence 2, Appl1
20	165.5	14.5	5889 4	US-09-310-235B-2	Sequence 2, Appl1
21	125	11.0	785 3	US-09-402-929-3	Sequence 3, Appl1
22	125	11.0	785 3	US-09-008-979A-6	Sequence 3, Appl1
23	125	11.0	785 4	US-09-460-618-6	Sequence 6, Appl1
24	119	10.5	2837 4	US-09-310-235B-6	Sequence 6, Appl1
25	105.5	9.3	1116 4	US-08-928-941D-17	Sequence 11, Appl1
26	105.5	9.3	1116 4	US-08-928-941D-17	Sequence 11, Appl1
27	105.5	9.3	2903 4	US-08-928-941D-2	Sequence 17, Appl1
28	105.5	9.3	2903 4	US-08-928-941D-3	Sequence 2, Appl1
29	105.5	9.3	2903 4	US-08-928-941D-3	Sequence 2, Appl1
30	105.5	9.3	2903 4	US-09-280-590A-2	Sequence 3, Appl1
31	105.5	9.3	3767 4	US-08-928-941D-28	Sequence 2, Appl1
32	105.5	9.3	3767 4	US-08-928-941D-30	Sequence 28, Appl1
33	105.5	9.3	3767 4	US-08-928-941D-30	Sequence 30, Appl1
34	105.5	9.3	3767 4	US-09-280-590A-28	Sequence 28, Appl1
35	96.5	8.5	282 2	US-08-814-030-2	Sequence 30, Appl1
36	93	8.2	336 1	US-08-248-474-105	Sequence 2, Appl1
37	93	8.2	336 1	US-08-248-474-105	Sequence 105, App
38	84.5	7.4	3834 4	US-08-756-849-105	Sequence 663, App
39	83.5	7.3	2581 4	US-09-370-838-66	Sequence 16, Appl1
40	82.5	7.2	3642 3	US-08-946-026-16	Sequence 23, Appl1
41	81.5	7.2	3627 1	US-08-447-500-23	Sequence 23, Appl1
42	81.5	7.2	3627 1	US-08-454-097-23	Sequence 23, Appl1
43	81.5	7.2	3627 1	US-08-453-866-23	Sequence 23, Appl1
44	81.5	7.2	3627 3	US-08-185-359-23	Sequence 23, Appl1
45	81.5	7.2	3628 1	US-08-468-036-17	Sequence 17, Appl1

ALIGNMENTS

RESULT 1
US-08-997-251-1
; Sequence 1, Application US/08997251
; Patent No. 6271440
GENERAL INFORMATION:
; APPLICANT: GUBLER, FRANZ J.
; APPLICANT: JACOBSEN, JOHN V.
; TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
; NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Boulevard
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,251
; FILING DATE: 23-DEC-1997
CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO A096/00383
; FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P6470/95
; FILING DATE: 09-NOV-1995
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P63779/95
; FILING DATE: 23-JUN-1995
ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P.

```

: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 110-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2220 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Hordeum vulgare
: STRAIN: Himalaya
: DEVELOPMENTAL STAGE: Seed
: TISSUE TYPE: Aleurone layer
: IMMEDIATE SOURCE:
: CLONE: HVCAMyb
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 275..1933
:
: US-08-997-251-1
:
: Alignment Scores:
: Pred. No.: 1,47e-38 Length: 2220
: Score: 396.00 Matches: 70
: Percent Similarity: 69.50% Conservative: 28
: Best Local Similarity: 49.65% Mismatches: 31
: Query Match: 34.80% Indels: 12
: DB: 4 Gaps: 1
:
: US-10-021-811-36 (1-206) x US-08-997-251-1 (1-2220)
:
: QY 13 ProgluValArgLysGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyrIle 32
: DB 392 CCGCCTGTGAAGAGGGGCGCTGGACGTCGCCGAGAGCAGCCATCTGCTGACTACGTG 451
:
: QY 33 AlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaLysLeuLysArgAsn 52
: DB 452 AAGAAAGCAGCGGAGGAGGAACTGGAAACGGGTGCAGAAAGAACACCGGCTTTCGGGTGC 511
:
: QY 53 GlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsn 72
: DB 512 GGCAAGAGCTGCCGCTCCGGTGGCGAACCACCTCAGGCCCAACCTCAAGAGGGGGCC 571
:
: QY 73 IleThrProGluGluGlnLeuIleuMetGluLeuHisAlaLysTrpGlyAsnArgTrp 92
: DB 572 TTCACCCCGAGAGAGAGAGCTCATCTCCAGCTCCACGATCGCAAGATGGCAAGATGG 631
:
: QY 93 SerLysIleAlaLysHisLeuProGlyArgTrpAspAsnGluIleLysAsnTyrTrpArg 112
: DB 632 GCTCGATGGCCGCTCATTTGCCAGGGCGTACTGATTAATAAAGAAATTAAGTCTGGAAC 691
:
: QY 113 ThrArgIleGlnLys-----HisIleLys 120
: DB 652 ACTCAATTAAGATGTCTACGAGCGGTTGCCAATATATCTGCTAGTGTATGCAT 751
:
: QY 121 GlnAlaGluAspHleGlnGlnIleSerSerAsnAsnSerGluIleAsnAspHisGlnAla 140
: DB 752 CAATCTCAAAATGAAGATCAGCAGGGCTCCAGCATTTCACTGCGGAGAAATCTTCC 811
:
: QY 141 Ser 141
: DB 812 AGT 814
:
: RESULT 2
: US-08-997-251-3
: Sequence 3, Application US/08997251
: Patent No. 6271440
: GENERAL INFORMATION:
: APPLICANT: GUBLER, FRANZ J.
: APPLICANT: JACOBSEN, JOHN V.
:
: TITLE OF INVENTION: PLANT REGULATORY PROTEINS III
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
: STREET: 5370 Manhattan Boulevard
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/997,251
: FILING DATE: 23-DEC-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO AU96/00383
: FILING DATE: 21-JUN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PN6470/95
: FILING DATE: 09-NOV-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PN3779/95
: FILING DATE: 23-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Winner, Ellen P.
: REGISTRATION NUMBER: 28,547
: REFERENCE/DOCKET NUMBER: 110-97
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 499-8080
: TELEFAX: (303) 499-8089
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2352 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 396..2054
:
: US-08-997-251-3
:
: Alignment Scores:
: Pred. No.: 2,69e-37 Length: 2352
: Score: 386.00 Matches: 68
: Percent Similarity: 69.57% Conservative: 28
: Best Local Similarity: 49.28% Mismatches: 30
: Query Match: 33.92% Indels: 12
: DB: 4 Gaps: 1
:
: US-10-021-811-36 (1-206) x US-08-997-251-3 (1-2352)
:
: QY 13 ProgluValArgLysGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyrIle 32
: DB 507 CCCCCGTGAAGAGGGGCCATGACGTGCGCGAGAGCGCCATCTGCTGACTACGTG 566
:
: QY 33 AlaAsnHisGlyGluGlyValTrpAsnSerLeuAlaLysAlaLysLeuLysArgAsn 52
: DB 567 AAGAAAGCAGCGGAGGAGGAACTGGAAACGGGTGCAGAAAGAACACCGGCTTTCGGGTGC 626
:
: QY 53 GlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsn 72
: DB 627 GGCAAGAGCTGCCGCTCCGGTGGCGAACCACCTGAGGCCCAACCTCAAGAGGGGGCC 686
:
: QY 73 IleThrProGluGluGlnLeuIleuMetGluLeuHisAlaLysTrpGlyAsnArgTrp 92
: DB 687 TTCACCCCGAGAGAGAGAGCTCATCTCCAGCTCCACGATGGGAGAACAGTGG 746
```

```

QY 93 SerIysIleAlaIySHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArg 112
Db 747 GCTGGATGGCGCTCATTTGCGAGGCGACGTATATGAATAAAGAAATTACTGCAAT 806
QY 113 ThrArgIleGlnLys-----HisIleLys 120
Db 807 ACTCGAATAAAGACATGCCAGCGAGCGCTACCCATCTATCTTACCAGCGTATGCAAT 866
QY 121 GlnAlaGluAsnPhenGlnGlnIleSerAsnAsnSerGluIleAsnAspHis 138
Db 867 CAATCTCAATGAAGATCAACAGCTGCTCCAGTATTTTGACTGTGGCGAAGAT 920

RESULT 3
US-08-722-626B-1
Sequence 1, Application US/08722626B
Patent No. 5939601
GENERAL INFORMATION:
APPLICANT: Yang, Yinnong
APPLICANT: Klessig, Daniel, F.
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED
TITLE OF INVENTION: DISEASE RESISTANCE IN PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/722,626B
FILING DATE: 27-SEP-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pat Hagan
REGISTRATION NUMBER: 27,643
REFERENCE/DOCKET NUMBER: 97-0010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215 563-4100
TELEFAX: 215 563-4044
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1344 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOETHERICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 148...981
OTHER INFORMATION:
US-08-722-626B-1

Alignment Scores:
Pred. No.: 3,09e-37 Length: 1344
Score: 382.50 Matches: 88
Percent Similarity: 54.38% Conservative: 30
Best Local Similarity: 40.55% Mismatches: 60
Query Match: 33.61% Indels: 39

```

```

DB: 2 Gaps: 6
US-10-021-811-36 (1-206) x US-08-722-626B-1 (1-1344)
QY 15 ValArgLysGlyProIleThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsn 34
Db 181 CTGAAAAAAGGCGCATGATTCCTGTGAAGAGATCAAGATTCCTCTTCATTCCAACT 240
QY 35 HisGlyGluGlyValTyrAsnSerLeuAlaLysAlaAlaGlyLeuLysArgGlnGlyLys 54
Db 241 AATGGCGATGGCAACGCGGAGCGCTCCCAACAGCGTGAGATATGGATGCGGGAG 300
QY 55 SerCysArgLeuArgThrLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThr 74
Db 301 AGTTGCGAGCTGGGTGGAGCATTTATTTGGCGACAGATATAAGAGGGGAATTTACCC 360
QY 75 ProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyLysAsnArgTrpSerLys 94
Db 361 AAGGAAGAGAACAAACAAATATCCAGTTACATGAATGAAATGCTTGCAATGATGCTGCA 420
QY 95 IleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArg 114
Db 421 ATAGCAGCAAAATTTACCGAGCAGACAGAACAAATGAATATAAATTTTGGCACACCCAC 480
QY 115 IleGlnLysHisIleLys-----GlnAla 122
Db 481 TTGAAGAGAGAGCTCAAGATTTAAAGCTCTCAGAACTCCAAAAGACATCCAAATGCC 540
QY 123 GluAsnPhenGln-----GlnIleSerSerAsnAsnSerGluIle--- 135
Db 541 AACAAATCATGATTCCAAGGCTCTACTACTCTGAAATCATCCAAATATTTCTATCTACT 600
QY 136 -----AsnAspHisGlnAla 140
Db 601 ATTATTAATACAAAAACAAACATTTAGCCAGTGTAGTCTTACCTCAACACCCAAATT 660
QY 141 SerThrSer-----HisValSerThrMetAlaGluProMetGluMetLysSerProCys 159
Db 661 TCATCTAGTACTGAATGATGCTCACTGTGACACTAGTCAGATGATCAATATGTTGTGATT 720
QY 160 TyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 179
Db 721 AACGAAGAAGTAATGAGAGTCG---TCCGAGTATTTTCCAGAGATCGAT-----GAG 768
QY 180 SerCysCysThrAsnAspAsnAsnAsnIleAsnTyrTrpSerMetGluAsp 196
Db 769 AGTTTGTGACGGAGCATTAACAAACGACAAATTAAGTGAAGTGAAGTGAAGTGAAGT 819

RESULT 4
PCT-US93-06251-90
Sequence 90, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCUDLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

```

NAME: Digilio, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8586
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 90:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1035 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 PCT-0593-06251-90

Alignment Scores:
 Pred. No.: 5,58e-28 Length: 1035
 Score: 305.50 Matches: 70
 Percent Similarity: 59.34% Conservative: 38
 Best Local Similarity: 38.46% Mismatches: 59
 Query Match: 26.85% Indels: 15
 Gaps: 7

US-10-021-811-36 (1-206) x PCT-US93-06251-90 (1-1035)

QY 12 AspProGluValArgLysGlyProThrPthMetGluGluAspLeuIleLeuMetAsnTyr 31
 DB 121 AACCTGAGCTCATCAAGAGGCTCTTGACCAAGAGAGATCAAGAGTATAGACTT 180
 QY 32 IleAlaAsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAlaGlyLeuLys--- 50
 DB 181 GTACGAAATACGGTCCCAACGGTGTCTGTATTGGCAAG-----CACTTAAAGGG 234
 QY 51 ArgAsnGlyLysSerCysArgLeuArgTyrPLeuAsnTyrIleuArgProAspValArgTyr 70
 DB 235 AGAATTGGAAATAATGAGGAGAGAGGATACCACTTGATTCAGAGAGTTAAAGAA 294
 QY 71 GlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyAsn 90
 DB 295 ACCTCTGTGACAGAGAGAGAGACAGAAATTATTACAGGACACACAGAGACTGGGAAAC 354
 QY 91 ArgTyrSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyr 110
 DB 355 AGATGGGAGAAATGCAAGACTACTGCTGAGCACTGATTAATGCTATCAAGAACAC 414
 QY 111 TPArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhenGlnGlnSerSer 130
 DB 415 TGGAAATTCACAAATGCTGCGAGAGTCCACAG--GAAGCTTATCTGACGAGCTTCA 471
 QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThMet 148
 DB 472 AAAGCCACCCAGCCAGCAGTGCACAAAGCTTCCAGAAAGACATTTGATGGGTTT 531
 QY 149 AlaGluProMetGluMetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSer 168
 DB 532 GCTCAG-----GCTCGCCCTACAGCTCA-----CTCCCTGCCACT 567
 QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTyrAsnAspAsnAsn 188
 DB 568 GCCAG-----CCACTGTTAACAACAGACTATTCATTCACAAATTTGGAAGCAAAAT 624
 QY 189 IleAsn 190
 DB 625 GTCTCC 630

RESULT 5
 US-08-306-691B-45
 Sequence 45, Application US/08306691B
 Patent No. 5734039
 GENERAL INFORMATION:
 APPLICANT: Calabretta, Bruno
 APPLICANT: Skorski, Tomasz

TITLE OF INVENTION: ANTISENSE
 TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
 NUMBER OF SEQUENCES: 55
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P. C.
 STREET: Two Penn Center, Suite 1800
 CITY: Philadelphia
 STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 720 KB
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WordPerfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/306,691B
 FILING DATE: September 15, 1994
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Monaco, Daniel A.
 REGISTRATION NUMBER: 30,480
 REFERENCE/DOCKET NUMBER: 8321-8
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-8383
 TELEFAX: (215) 568-5549
 TELEX: NO. 5734039e
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3225 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-306-691B-45

Alignment Scores:
 Pred. No.: 3.13e-27 Length: 3225
 Score: 305.50 Matches: 70
 Percent Similarity: 59.34% Conservative: 38
 Best Local Similarity: 38.46% Mismatches: 59
 Query Match: 26.85% Indels: 15
 Gaps: 7

US-10-021-811-36 (1-206) x US-08-306-691B-45 (1-3225)

QY 12 AspProGluValArgLysGlyProThrPthMetGluGluAspLeuIleLeuMetAsnTyr 31
 DB 372 AACCTGAGCTCATCAAGGCTCTTGACCAAGAGAGATCAAGAGTATAGACTT 431
 QY 32 IleAlaAsnHisGlyGluGlyValTTPAsnSerLeuAlaLysAlaGlyLeuLys--- 50
 DB 432 GTACGAAATACGGTCCCAACGGTGTCTGTATTGGCAAG-----CACTTAAAGGG 485
 QY 51 ArgAsnGlyLysSerCysArgLeuArgTyrPLeuAsnTyrIleuArgProAspValArgTyr 70
 DB 486 AGAATTGGAAATAATGAGGAGAGTGCATTAACCACTTGATTCAGAGAGTTAAAGAA 545
 QY 71 GlyAsnIleThrProGluGluGlnLeuLeuIleMetGluLeuHisAlaLysTyrGlyAsn 90
 DB 546 ACCTCTGTGACAGAGAGAGACAGAAATTATTACAGGACACACAGAGACTGGGAAAC 605
 QY 91 ArgTyrSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyr 110
 DB 606 AGATGGGAGAAATCCCAAGACTACTGCTGAGCAGAGATGATTAATCTTCAAGAACAC 665
 QY 111 TPArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhenGlnGlnSerSer 130
 DB 666 TGGAAATTCACAAATGCTGCGAGAGTCCACAG--GAAGCTTATCTGACGAGCTTCA 722
 QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThMet 148

Db 723 AAAGCCAGCCAGCCAGGAGGCGCCACAAAGCTTCACGAGAAAGACTCATTTGATGGGTTT 782
QY 149 AAlaGluProMetGluMetTySerProCysTyrgInGlyMetLeuGluProPheSer 168
Db 783 GCTCAG-----GCTCGGCTACAGCTCAA-----CTCCCTGCCACT 818
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188
Db 819 GGCCAG---CCCACTGTTAACAAGCACTATTCCTATTACCAACATTTCTGAAGCAAAAT 875
QY 189 IleAsn 190
Db 876 GTCTCC 881
RESULT 6
PCT-US93-06251-91
Sequence 91, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 3225 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-91
Alignment Scores:
Pred. No.: 3,13e-27 Length: 3225
Score: 305.50 Matches: 70
Percent Similarity: 59.34% Conservative: 38
Best Local Similarity: 38.46% Mismatches: 59
Query Match: 26.85% Indels: 15
Gaps: 7
US-10-021-811-36 (1-206) x PCT-US93-06251-91 (1-3225)
QY 12 AspProGluValAlaGlySerGlyProTPrhMeGluGlnAspLeuIleuMetAsnTyR 31
Db 372 AACCTTGAGCTCAATCAAGGCTCTTGACCAAGAAAGATGAGAGATGAGACTT 431
QY 32 IleAsnAsnHisGlyGluGlyValTyPAsnSerLeuAlaIleAlaGlyLeuLys--- 50
Db 432 GTACAGAAATAAGGTCGAAAGCTGTGCTGTTATTGCCAAG-----CACTTAAAGGG 485

QY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyRLeuArgProAspValArgArg 70
Db 486 AGAATTGGAAACAAATGATGAGGAGAGGTGCATTAACCACTTGATTCACAAATTAAGAA 545
QY 71 GlyAsnIleThrProGluGluGlnLeuIleMetGluLeuHisAlaLysTrpGlyAsn 90
Db 546 ACCTCGTGCACAGAAAGACAGAAATTAATTACACAGCACAGAAAGCTGGGAAAC 605
QY 91 ArgTrpSerLysIleAlaLysHisLeuProGlyArgTrpThrAspAsnGluIleLysAsnTyR 110
Db 606 AGATGGCAGAAATCCCAAGACTACTGCTGGACGAATGATTAATCTCAAGAACAC 665
QY 111 TrpArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhenGlnInsSerSer 130
Db 666 TGAATTCTACAATGCGTCGGAAGGTCAAGACAG---GAAGCTTACTCGAGAGCTTCA 722
QY 131 AsnAsnSerGlu-----IleAsnAspHisGlnAlaSerThrSerHisValSerThreMet 148
Db 723 AAAGCCAGCCAGCCAGCACTGGCCACACAGCTTCAGAAAGCAAGCATTTGATGGGTTT 782
QY 149 AAlaGluProMetGluMetTySerProCysTyrgInGlyMetLeuGluProPheSer 168
Db 783 GCTCAG-----GCTCGGCTACAGCTCAA-----CTCCCTGCCACT 818
QY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn 188
Db 819 GGCCAG---CCCACTGTTAACAAGCACTATTCCTATTACCAACATTTCTGAACACAAAT 875
QY 189 IleAsn 190
Db 876 GTCTCC 881
RESULT 7
PCT-US93-06251-89
Sequence 89, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 3230 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-89

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3.14e-27	305.50	59.34%	38.46%	26.85%	3230	70	38	59	15	7

US-10-021-811-36 (1-206) x PCT-US93-06251-89 (1-3230)

OY 12 AspProgluValArgylsGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31
 Db 262 AACCTGAGCTCATCAGAGGTCCTGGACCAAGAGATCATGAGAGTATGAGACTT 321
 OY 32 IleAlaAsnHISGLYLGLYLGLYAlTrpAsnSerIleuAlaLysAlaLysLysLeu--- 50
 Db 322 GTACGAAATACGTCGCCAAGCGTTGGCTCTTATGGCAAG-----CACTTAAAGGG 375
 OY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArg 70
 Db 376 AGAATTGGGAAACATGTAGGAGAGGTGGCATACCACTTGAATCCAGAAATTAGAA 435
 OY 71 GlyAsnIleThrProGluGluGluInleuIleuMetGluLeuHISAlaLysTrpGlyAsn 90
 Db 436 AACCTCTGGACAGAGAGAGAGATTTATTCAGGACACAGAGAGCTGGGAGAC 495
 OY 91 ArgTrpSerLysIleAlaLysHISLeuProGlyArgTrpAsnGluIleLysAsnTyr 110
 Db 496 AGATGGGAGAAATCGCAAACTGCTGCTGAGCAAGTGTATGCTATCAGAAACCC 555
 OY 111 TrpArgTrpArgIleGluLysHISLysGluInleuAsnPhenGluInleuSerSer 130
 Db 556 TGGATTTCTACAAATCGTCGAGAGGTGCAGACAG--GAAGTTATTCGACAGATCTTCA 612
 OY 131 AsnAsnSerGlu-----IleAsnAspHISGlnAlaSerTrpSerHISValSerThrMet 148
 Db 613 AAAGCCAGCCAGCCAGCGAGCCACAAAGCTTCCAGAAAGACATTTGATGGTTT 672
 OY 149 AlaGluProMetGluMetLysSerProProCysTyrGlnGlyMetLeuGluProPheser 168
 Db 673 GTCAG-----GCTCGGCTTACAGCTCA-----CTCCCGCCACT 708
 OY 169 ThrGlnPheProThrIleAsnProAspGlnSerSerCysTrpAsnAspAsnAsn 188
 Db 709 GCCCGAG--CCCATGTTAACACACACATTCATTTACACATTTTGAAGCACAAAT 765
 OY 189 IleAsn 190
 Db 766 GTCCTCC 771

RESURF 8
 US-09-167-322-10
 ; Sequence 10, Application US/09167322
 ; Patent No. 6365151
 ; GENERAL INFORMATION:
 ; APPLICANT: Allegheny University of the Health
 ; Sciences, Halpern, Michael S.
 ; England, James M.
 ; TITLE OF INVENTION: CANCER VACCINE
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
 ; STREET: Suite 1800, Two Penn Center Plaza
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19102
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/167, 322
 FILING DATE: 07-Oct-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/00582
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Monaco, Daniel A.
 REGISTRATION NUMBER: 30,480
 REFERENCE/DOCKET NUMBER: 7933-33 PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-8383
 TELEFAX: (215) 568-5549
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1897 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:
 US-09-167-322-10

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2.04e-26	296.00	53.47%	36.14%	26.01%	1897	73	35	55	40	9

US-10-021-811-36 (1-206) x US-09-167-322-10 (1-1897)

OY 12 AspProgluValArgylsGlyProTrpThrMetGluGluAspLeuIleuMetAsnTyr 31
 Db 233 AACCTGAGCTCATCAGAGGTCCTGGACCAAGAGATCATGAGAGTATGAGACTT 292
 OY 32 IleAlaAsnHISGLYLGLYLGLYAlTrpAsnSerIleuAlaLysAlaLysLysLeu--- 50
 Db 293 GTCCGAAATACCGTCCAAAGCGCTGTCGACATTTGCTAAG-----CATTTGAAGGA 346
 OY 51 ArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeuArgProAspValArgArg 70
 Db 347 AGATTTGGAAACATGTAGGAGAGGTGCACACACATGTATGATTCAGAAAGTGA 406
 OY 71 GlyAsnIleThrProGluGluGluInleuIleuMetGluLeuHISAlaLysTrpGlyAsn 90
 Db 407 AACCTCTGGACAGAGAGAGAGATTAATTTATTCAGGACACACAGAGACTGGGAAAC 466
 OY 91 ArgTrpSerLysIleAlaLysHISLeuProGlyArgTrpAspAsnGluIleLysAsnTyr 110
 Db 467 AGATGGGAGAAATTTGCAAAAGTTGCTGCTGGACGAGATGATACGCTGTCAAGAAC 526
 OY 111 TrpArgThrArgIleGluLysHISLysGluAlaGluAsnPhenGluInleuSerSer 130
 Db 527 TGGAAATTCACATGCGCGGAGAGGTGCAGAG--GAGGTTATCCCGAGAGTCTCC 583
 OY 131 Asn-----AsnSerGluIleAsnAspHISGlnAlaSerThrSerHISValSerThr 147
 Db 584 AAAGCGGGCGCGGCTCGGACACACCGGCTTCCAAAGAC--AGCATCTGATGGC 640
 OY 148 MetAlaGluProMetGluMetLysSerProProCys----- 159
 Db 641 TTTGCC-----CACAAACCACTGCGAGGCCGCTCCCGGGGCGGCCAG 685
 OY 160 -----TyrGlnGlyMetLeuGluProPheserThr----- 169
 Db 686 GCCCTCTGGGACGATGACTACCTTACTACACATTTGAGGCAAAATGCTCCCTGCT 745
 OY 170 GlnPhePro-----ThrIleAsnProAspGlnSerSerCys 181
 Db 746 CAGATCCCATTCACAGTACGACTGATATAATATATATATCAAGTTCCTCAGCC-ACGTC 804

QY 182 CysThr 183
DB 805 TGCACG 810

RESULT 9
US-09-402-929-5
Sequence 5, Application US/09402929
Patent No. 6410825

GENERAL INFORMATION:
APPLICANT: Temple University - Of The Commonwealth System of Higher Education
APPLICANT: Toscani, Antonio
APPLICANT: Hatton, Kimi
APPLICANT: Reddy, E. P.
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/06896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-214 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4880 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-402-929-5

Alignment Scores:
Pred. No.: 2e-25 Length: 4880
Score: 293.00 Matches: 70
Percent Similarity: 56.22% Conservative: 34
Best Local Similarity: 37.84% Mismatches: 57
Query Match: 25.75% Indels: 24
Gaps: 8

US-10-021-811-36 (1-206) x US-09-402-929-5 (1-4880)

QY 12 AspprogluValArglySGlyProtrPTThMeGluGluaspLeuileuMeIAsnTyr 31
DB 348 AATCTTAATTTGATTAAGGCTCTTGACCTAAAGACAGATCAGACGGTTATTGAATTA 407

QY 32 lLeAlaasnHsIsglyGluGlyValTTPAsnSerIeuAlaIysAlaIagIleuLys--- 50
DB 408 GTTCAGAAATATGGCCAAAAGATGCTTTTATTGCAAAA-----CATTTAAAAAGA 461

QY 51 ArgSsnIySserCysArgLeuArgTTPLeuAsnTyrIleuArgProAspValAArgarg 70
DB 462 AGAATATAGCAAGCAGTGTAAAGAAAGATGACATATCATCTGAAATCTGAGTAAAGAAA 521

QY 71 GlyAsnIleThrProGluGluInLeuLeuIleuMeGluLeuHsAlaIysTTPGlyAsn 90

DB 522 TCCTCGGACAGAAAGAGAGACAGCATCATCTATGAAAGACATAGCGGTTGGAAAT 581

QY 91 ArgTTPserLysIleAlaIysHsIleuProGlyArgTTPAspSngIleuLysAsnTyr 110
DB 582 CGTTGGCAGAAATTTGCCAAACTACTTCCAGAGAGAGATGATATCTATCAAAAATCAT 641

QY 111 TTPArgThrArgIleGlnLysHsIleYsgIAlaGluAsnPhgIngIn-----Gln 128
DB 642 TGGAAATCTACTATGCCAAGAAAGTGAAACAGAGAGCGCATTTTACAAGATGAATMAA 701

QY 129 SerSerAsnSngSerGluIleAsnAspHisGln-----AlaSerThrSerHisValSer 146
DB 702 TCAGAACGATCTTCACTTAACCTTCAACACAAACCTTGACACTATGATCATATGCAA 761

QY 147 ThrMetAlaGluProMetGluMetTyrSerPro-----ProCysTyrGlnIlyMet 163
DB 762 ACCCAG-----AATCAGTTTACATACCTGTTACAGATCCCTGGATACGATATGTG 812

QY 164 -----LeuGluProPhSerThr-----GlnPhePro 172
DB 813 TCACCTGAAGCAATGTATAGAACATGTTACAGCTTACTTCCCTTTATTCAGCAACC 872

QY 173 ThrIleAsnProAsp 177
DB 873 TTCATTGATGATGAT 887

RESULT 10
US-09-402-929-1
Sequence 1, Application US/09402929
Patent No. 6410825

GENERAL INFORMATION:
APPLICANT: Temple University - Of The Commonwealth System of Higher Education
APPLICANT: Toscani, Antonio
APPLICANT: Hatton, Kimi
APPLICANT: Reddy, E. P.
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/06896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 6056-214 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3602 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-402-929-1

Alignment Scores:	
Pred. No.:	1,92e-25
Score:	291.50
Percent Similarity:	59.49%
Best local Similarity:	40.51%
Query Match:	25.62%
DB:	4
Length:	3602
Matches:	64
Conservative:	30
Mismatches:	51
Indels:	13
Gaps:	6

Oy	12	AspProGluValaLgysglYpRtPThrIleMetIuLyspLeuIleLeuMetAsnTyr	31
Db	500	AATCCAGAAATGCATTAAGAAGCGCTCTGGACATGAAGAAAGATCAGAGGGTTATTGCAATTA	555
Oy	32	IleAlaAsnHisIsgLylgLuIglYValITrpAsnSerLeuAlaLysAlaAlaGlyLeuLys---	50
Db	560	GTTCAGAAATTTGGGCCCAAAAAGGTGGCTTTTATTCGCAAAA-----CATTTTAAAGGA	613
Oy	51	ArgAsnGlyLysSerCysArgLeuArgITrpLeuAsnTyrLeuArgProAspValaArg	70
Db	614	AGAAATAGCCAAAGCGAGTGCAGAGAAAGATGCAGCAATACCTGAAACCTGAAAGTGAAGAAG	673
Oy	71	GlyAsnIleThrProIgluIgluIgluLeuLeuIleMetIuLeuHisAlaLysTyrGlyLysn	90
Db	674	TCTTCCTGAGACAGAAAGAAAGACAGAGATCATATATGAGACACAAACCGCGCTGGAAAC	733
Oy	91	ArgITrSerLysIleAlaLysHisLeuProLysArgIThrLysAsnGlyIleLysAsnTyr	110
Db	734	CGTTGGGCGCGAGATTGCTAGATCTACTCTCTCGGAAGAGCTGATAATTTCTATCAAAAAATCAT	793
Oy	111	TrpArgIThrArgIleGlnLysHisIleLysGlnAlaLysAsnPheGlnGln-----Gln	128
Db	794	TGGAATTCCTACCAGTCGCAAGAAAGATGGAAACAGAGGGCTATTTCACAAAGATGCAATAAAA	855
Oy	129	SerSerAsnAsnSerGluIleAsnAspHisGln-----AlaSerIThrSerHisValSer	146
Db	854	TCAAGACGCGCTTCATCAAAACCTTCAACACACAAACCTTGGAGACTAATAGAACCAATTTCCAA	913
Oy	147	ThrMetAlaGluPromeGlnMetLysSerPro-----ProCysTyrGln	161
Db	914	ACCCAG-----AATCACTTTTACATTCCTCTGTACATCCCTGGGATACG	958

```

?      ATTORNEY/AGENT INFORMATION:
?      NAME: Monaco, Daniel A.
?      REGISTRATION NUMBER: 30,480
?      REFERENCE/DOCKET NUMBER: 6056-214
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (215) 568-8383
?      TELEFAX: (215) 568-5549
?      INFORMATION FOR SEQ ID NO: 4:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 6775 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
US-09-402-929-4

```

```

Oy 12 AsprGtGValaLagLySgCyLProTPrHMeGluGlaSLeuIleuMetKsnYr 31
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 3749 AATCCAGATATGATAAAGGGCTCTTGAGCTAAGGAGGAAGATCAGAGGTTTGAATTA 3808
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 32 ILlaLsnhISgLyGluGlyValITrPaSnSerLeuAlaLyAlaIaGlyLeuLys--- 50
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 3809 GTTTCAGAAATATGGGCCCAAAAAGCGCTCTTAAATGGCAAAA-----CATTTAAAGCA 3862
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 51 ArgAsnGlyLysSerCysArgLeuAcrGTrPLeuAsnTyLeuArgProAspValaArg 70
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 3863 AGAATAGGCAAGCAGCGCTCAAGAAAGAATGGCAGCATCACCCTTAACCCGTGAAGTGAAGAG 3922
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 71 GlyAsnIleThrProGluGluGluLeuLeuIleMetGluLeuHisAlaLysTrpCysAsn 90
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 3923 TCTTCTCGGACAGAGAAAGAAAGACAGAGCATATGAAAGCACACAAAGCGCTGGGAAC 3982
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 91 ArgTrpSerLysIleAlaLysHisLysLeuProGlyArgTrpAspAsnGluIleLysAspTr 110
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 3983 CGTTGGGCGCGAGATGCTAAAGTACTTCCGTGGAGAGAGCTGAATTAATCTTATCAAAATATAT 4042
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 111 TrpArgTrhArgIleGluLysHisLileLysGlnAlaGlaAspPheGlnGln-----Gln 128
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 4043 TGGAAATTCATCCACAGCGCAAGAAAAAGTGAACACAGAGGCGTATTTCACAGATGCAATAAAA 4102
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 129 SerSerAsnAsnSerGluIleAsnAspHisGln-----AlaSerTrpSerHisValSer 146
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 4103 TCAAGAGGCGGCTTCATCAAAAACCTTCAACAAACCTTGGCAGCATATGAGCATTATGCA 4162
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Oy 147 ThrMetAlaGluProMetGluMetGlySerPro-----ProCysTrpGln 161
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 4163 ACCCAG-----AATCAGTTTACATTCCTGTTTCAAGATCCCTGGGATATCAG 4207
    :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::

```

```

COMPUTER READABLE FORM :
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/06896
FILING DATE:

```

RESULT 11
US-09-402-929-4
Sequence 4, Application US/09402929
Patent No. 6410825
GENERAL INFORMATION:
APPLICANT: Temple University - Of The Commonwealth System of Higher Education
APPLICANT: Toscani, Antonio
APPLICANT: Hatton, Kimi
APPLICANT: Reddy, E. P.
TITLE OF INVENTION: A-myb NULL MUTANT TRANSGENIC ANIMALS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/402,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/06896
FILING DATE:

ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,691B
FILING DATE: September 15, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: NO. 5734039e
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 2638 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-306-691B-46

Alignment Scores:
Pred. No.: 3 37e-23 Length: 2638
Score: 271.50 Matches: 67
Percent Similarity: 52.17% Conservative: 29
Best Local Similarity: 36.41% Mismatches: 59
Query Match: 23.86% Gaps: 29
DB: 1 8
US-10-021-811-36 (1-206) x US-08-306-691B-46 (1-2638)
QY 5 GlnGlnCysLys-----ThrSerGlnAspProGluValAlaGlySGlyProTyr 20
Db 326 CAGCAATGCCAGTGCAGGTGGCTGAGAGTTTGTGAATCCAGACCTTCGCAAGGGCCCTGC 365
QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTyrTlleAlaAsnHisGlyGluGlyValTyr 40
Db 386 ACCAAAGAGGAGACCAAAAGATCATGACCTGTTAAGAGTATGCGACAAAGCAAGTGG 445
QY 41 AsnSerLeuAlaLysAlaIleGlyLeuLys--ArgAsnGlyLysSerCysArgLeuArg 59
Db 446 ACACTGATTCGCAAG-----CACCTAAGGCCGCCGCTGGGGAAGCAGTCCCTGAAACGC 499
QY 60 TrpLeuAsnTyrLeuArgProAspValAlaArgGlyAsnIleThrProGluGluGlnLeu 79
Db 500 TGGCACAACCACTCAACCTGAGTGAAGACAGTCTGCTGACCGCAGAGCGAGAGACCCG 559
QY 80 LeuIleMetGluLeuHisAlaLysTyrGlyAsnArgTyrSerLysIleAlaLysHisLeu 99
Db 560 ATCTATCTGCGAGGCCCAAGAGTCTGGGCAACCGCGCGCCAGATCGCCAAAGATGTTG 619
QY 100 ProGlyArgThrAspAsnGluIleLysAsnTyrTyrParGlyThrArgIleGlnLysHisIle 119
Db 620 CCAAGGAGGAGACAGATGCTGTGAACAATCAGTCACTGCACTCATCAAAAGAAAGTGTG 679
QY 120 LysGlnAlaLysAsnProGluGlnGlnSerSerAsnAsnSer----- 133
Db 680 ---GACACAGAGAGCTTCTTGAGCGATGCCAAGACTGCAAGCCCAAGTACTGCTGCTG 736
QY 134 ---GluIleAsnAspHisGlnAlaIleSerThrSerHisValSerThrMetAlaGluProMet 152
Db 737 CTGAGAGCTCGAGAGCAAGAGCGGCTTCACAGT-----GCCACGCCACAG 781
QY 153 GluMetLysSerProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhePro 172

Db 782 GAA-----GGCAGGAGGAAGTCT-----CTGACCACTGCGCC 814
QY 173 ThrIleAsnPro 176
Db 815 TCCGTCCTCCT 826

RESULT 13
US-08-485-139-1
Sequence 1, Application US/08485139
Patent No. 5880331
GENERAL INFORMATION:
APPLICANT: KREBBERS, Enno
APPLICANT: WILLIAMS, Mark
APPLICANT: LEEHMAN, Jan
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN
TITLE OF INVENTION: MALE STERILE PLANTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,139
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-096
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4059 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: C1 gene of Zea mays
FEATURE:
NAME/KEY: -
LOCATION: 279..284
OTHER INFORMATION: /label= HpaI
FEATURE:
NAME/KEY: -
LOCATION: 447..452
OTHER INFORMATION: /label= EcoRI
FEATURE:
NAME/KEY: -
LOCATION: 1735..1740
OTHER INFORMATION: /label= AatII
FEATURE:
NAME/KEY: -
LOCATION: 1505..1510
OTHER INFORMATION: /label= EcoRI
FEATURE:
NAME/KEY: -
LOCATION: 2081..2086
OTHER INFORMATION: /label= XhoI
FEATURE:
NAME/KEY: -
LOCATION: 2418..2430
OTHER INFORMATION: /label= SfiI

LOCATION: 447..452
OTHER INFORMATION: /label= ECORI
FEATURE:
NAME/KEY: -
LOCATION: 1735..1740
OTHER INFORMATION: /label= AatII
FEATURE:
NAME/KEY: -
LOCATION: 1505..1510
OTHER INFORMATION: /label= ECORI
FEATURE:
NAME/KEY: -
LOCATION: 2081..2086
OTHER INFORMATION: /label= XhoI
FEATURE:
NAME/KEY: -
LOCATION: 2418..2430
OTHER INFORMATION: /label= SfiI
FEATURE:
NAME/KEY: -
LOCATION: 2669..2674
OTHER INFORMATION: /label= SnaBI
FEATURE:
NAME/KEY: -
LOCATION: 2634..2639
OTHER INFORMATION: /label= SnaBI
FEATURE:
NAME/KEY: -
LOCATION: 3008..3013
OTHER INFORMATION: /label= HpaI
FEATURE:
NAME/KEY: -
LOCATION: 1..1077
OTHER INFORMATION: /label= PC1
OTHER INFORMATION: /note= "region containing promoter of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 1078..2134
OTHER INFORMATION: /label= C1
OTHER INFORMATION: /note= "coding region of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 2135..2430
OTHER INFORMATION: /label= 3'C1
OTHER INFORMATION: /note= "region containing polyadenylation signal of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 1033..1038
OTHER INFORMATION: /label= TATA-Box
FEATURE:
NAME/KEY: -
LOCATION: 1061..1062
OTHER INFORMATION: /label= transcript-init
OTHER INFORMATION: /note= "transcription initiation site"
FEATURE:
NAME/KEY: intron
LOCATION: 1211..1299
FEATURE:
NAME/KEY: intron
LOCATION: 1430..1575
FEATURE:
NAME/KEY: -
LOCATION: 935..939
OTHER INFORMATION: /label= C1-S
OTHER INFORMATION: /note= "GCAG sequence (in C1 gene) which in the C1-S sequence
OTHER INFORMATION: is changed to TTAGG"
US-08-750-357-1
Alignment Scores:
Pred. No.: 4,89e-20 Length: 4059
Score: 248.00 Matches: 62
Percent Similarity: 41.67% Conservative: 13
Best Local Similarity: 34.44% Mismatches: 26

Query Match: 21.79% Indels: 81
DB: 3 Gaps: 2
US-10-021-811-36 (1-206) x US-08-750-357-1 (1-4059)
QY 15 ValArgLyGlyProThrMetGluAspLeuIleuMetAsnTrpIleuAsn 34
DB 1111 GTTAAAGAGAGGGCGCGAGCAGCAGAGAGAGAGATGCTTGCGCGCTACGTCAAGGCC 1170
QY 35 HSGLyGluGlyValTrpAsnSerLeuAlaLys----- 45
DB 1171 CATGCCAAGGCAAAATGAGGGAAGT-GCCCCAGAAAGCCGTTAAACTAGTACTGTTT 1229
QY 45 ----- 45
DB 1230 TTATTTGATTTTGGCATCATATATATACCCCGAGGCAAGACCGAGACGATCACGTGT 1289
QY 46 ---AlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuAlaTrpLeuAsnTrpLeu 64
DB 1290 GTGGGTGCAGGTTTGGCTGCTGCGGCAAGAGCTGCCGCTCGGTGGCTGAATCACTC 1349
QY 65 ArgProAspValArgArgGlyAsnIleThrProGluGluLeuLeuIleMetIleu 84
DB 1350 CGGCCCAACATCAGCGCGGCAACATCTCTACGACGAGAGAGATCTCATCCGCTC 1409
QY 84 ----- 84
DB 1410 CACAGGCTCTCGGCAACAGCTGTGTCAGTGGCCAGTGGCTAGCTTATTACAGCA 1469
QY 84 ----- 84
DB 1470 GCTGACGACGAGCGCATCGATCGAGCGTCTGCGAAATCATCTGTCGGGTGCGGCC 1529
QY 85 -----HisAlaLysTrpGlyAsnArgTrpSerLysIle 95
DB 1530 GTGTGAGATGAGCTCATATCATATGACTGCTGTGGCG-CGCAGGTGGTGGCTGATT 1588
QY 96 AlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTrpArgThrArgIle 115
DB 1589 GCAGGAGGCTGCTGCGGCAACAGCATGAAATCAAGAACTACTGGAACGACGCTG 1648
RESULT 15
US-08-485-139-5
; Sequence 5, Application US/08485139
; Patent No. 5880331
; GENERAL INFORMATION:
; APPLICANT: KREBBERS, Enno
; APPLICANT: WILTIAMS, Mark
; TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN
; TITLE OF INVENTION: MALE STERILE PLANTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,139
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 010830-096
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4824 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: plasmid pCOL9
FEATURE:
NAME/KEY: -
LOCATION: 396..401
OTHER INFORMATION: /label= EcoRI
FEATURE:
NAME/KEY: -
LOCATION: 2367..2379
OTHER INFORMATION: /label= SfiI
FEATURE:
NAME/KEY: -
LOCATION: 884..888
OTHER INFORMATION: /label= C1-S
OTHER INFORMATION: /note= "TGCAG (in C1) which in C1-S allele is
replaced with TTAGG"
US-08-485-139-5

Alignment Scores:
Pred. No.: 6,35e-20 Length: 4824
Score: 248.00 Matches: 62
Percent Similarity: 41.67% Conservative: 13
Best Local Similarity: 34.44% Mismatches: 26
Query Match: 21.79% Indels: 81
DB: 2 Gaps: 2

US-10-021-811-36 (1-206) x US-08-485-139-5 (1-4824)

QY 15 ValArgLyGlyProTrpThrMetGluGluSpleuIleLeuMetAsnTyrIleAlaAsn 34
|||||:||||| |||||:||||| |||||:|||||
DB 1060 GTTAAGAGAGGGGCGCTGGACGACGACGACGATGCTTGCGCGCTACGTCAGAGCC 1119
QY 35 HisGlyGluGlyValTrpAsnSerLeuAlaLys----- 45
|||||:||||| |||||:||||| |||||:|||||
DB 1120 CATGGCAAGCAATGAGGAGAGT-GCCCGAAGCGGTAAACTAGTACTTT 1178
QY 45 ----- 45
DB 1179 TTATTTCATTGGGATCATATATACCCCGAGGCAACCGGAGAGATCAGTGT 1238
QY 46 --AlaAlaGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrpLeuAsnTyrLeu 64
|||||:||||| |||||:||||| |||||:||||| |||||:|||||
DB 1239 GTGGGTGCGAGTTTCGTCGCGCAAGACCTCCGCGTGGCTGACTACCTC 1298
QY 65 ArgProAspValArgArgGlyAsnIleThrProGluGluInLeuLeuIleMetGluLeu 84
|||||:||||| |||||:||||| |||||:||||| |||||:|||||
DB 1299 CGGCCCAACATCAGCGCGGCAACATCTCTACGAGAGAGATCTCATCCGCTC 1358
QY 84 ----- 84
DB 1359 CACAGGCTCTCGGCAACAGCTCTGTGAGTGGCCAGTGGGCTAGCTTATTACGA 1418
QY 84 ----- 84
DB 1419 GCTGACGACGAGCGATCGATCGACGCTGTCGCAATTCATCTCCGTGTGCGCC 1478
QY 85 -----HisAlaLysTrpGlyAsnArgTrpSerLysIle 95
|||||:||||| |||||:||||| |||||:||||| |||||:|||||
DB 1479 GTGTGAGAGTGAATTCATTCATATGTACATGGGTGTGGCG-CGCAAGTGGTGGCTGATT 1537
QY 96 AlaLysHisLeuProGlyArgThrAspAsnGluIleLysAsnTyrTrpArgThrArgIle 115
|||||:||||| |||||:||||| |||||:||||| |||||:|||||
DB 1538 GCAGCGAGGCTGCTGCGCGAAGACACATGAAATCAAGAACTACTGGAACAGCACGCTG 1597

Search completed: February 18, 2003, 11:44:39
Job time : 56 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 18, 2003, 09:25:50 ; Search time 220 Seconds
(without alignments)
2108.689 Million cell updates/sec

Title: US-10-021-811-36
Perfect score: 1138
Sequence: 1 MDRKQCKTSQDPEYRKGPW.....NNINYSMEDSMQNLNGD 206

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame-p2n.model -DEV-xmlh
-O-/cgu2.1/USPTO.spool/US10021811/runat.12022003.090133.9544/app.query.fasta.1.391
-DB-N.Geneseq.101002 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOCT=0
-DOOEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-BLOSUM62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=plc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -090133.9544 -NCPU=6 -ICPU=3
-NARN TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YCAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	759.5	66.7	862	21	AAC57336
2	671	59.0	910	21	AAC40587
3	667	58.6	908	21	AAC48506
4	625.5	55.0	907	21	AAC57337
5	585	51.4	621	21	AAC41714
6	585	51.4	1127	21	AAC37953
7	540	47.5	673	21	AAC56197
8	539	47.4	1137	22	AAD05784
9	523.5	46.0	981	24	ABK65252
10	510	44.8	972	24	ABK65364
11	495	43.5	626	22	AAH87724
12	478	42.0	524	22	AAC57194
13	475	41.7	837	20	AAC25572
14	459	40.3	959	21	AAC44894
15	421	37.0	389	21	AAC56152
16	421	37.0	417	21	AAC56774
17	410.5	36.1	516	21	AAC56352
18	410.5	36.1	516	21	AAC56457
19	406.5	35.7	1384	21	AAC42228
20	406	35.7	660	24	ABK65183
21	405.5	35.6	542	21	AAC56434
22	404.5	35.5	389	21	AAC56837
23	404	35.5	1081	22	AAFG0596
24	402.5	35.4	1820	21	AAZ46048
25	402.5	35.4	1820	22	AAH28336
26	400.5	35.2	774	21	AAC42662
27	397.5	34.9	1209	22	AAD05748
28	397.5	34.9	1209	24	ABK65181
29	396.5	34.8	1006	22	AAFG0594
30	396	34.8	2220	18	AAT60744
31	392	34.4	1218	21	AAC57346
32	391.5	34.4	612	24	ABK65317
33	390.5	34.3	800	24	ABK65247
34	390	34.0	526	21	AAC57245
35	387	34.0	1369	22	AAD05764
36	386.5	34.0	1825	21	AAC44207
37	386	33.9	2032	18	AAT60745
38	385	33.8	1062	21	AAC43611
39	385	33.8	1087	22	AAD06456
40	384	33.7	1081	24	ABK65248
41	383.5	33.7	949	22	AAFG5188
42	383.5	33.7	2521	20	AAZ55571
43	383	33.7	573	21	AAC57325
44	382.5	33.6	843	21	AAC43054
45	382.5	33.6	1044	24	ABN98399

ALIGNMENTS

RESULT 1
AAC57336
ID AAC57336 standard; DNA; 862 BP.

AC AAC57336;
XX 25-JAN-2001 (first entry)
XX
XX
XX
XX

Eucalyptus grandis transcription factor DNA sequence #773.

Platt; transcription factor; gene expression; eucalyptus; pine; acacia;
poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;
type 2 Cys2His2; CCAAT box element; MYB; ss.

OS Eucalyptus grandis.
XX

PN WO200053724-A2.
 XX 14-SEP-2000.
 XX 09-MAR-2000; 2000WO-US06112.
 PF 11-MAR-1999; 99US-0266513.
 PR 18-AUG-1999; 99US-0149485.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX Wood M, McGrath A, Shenk MA, Glenn M;
 XX WPI. 2000-579369/54.
 DR
 XX
 PT New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT
 XX
 PS Claim 1; Page 649; 747pp; English.
 XX
 CC The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
 CC sequence for one such transcription factor. The transcription factor may
 CC be used to produce a plant having modified gene expression such as a
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
 CC mahogany species or to modify the activity of a polypeptide in a plant.
 CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bzip, bzip family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC homeobox/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
 CC and ERBBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and XTB.
 XX
 SO Sequence 862 BP; 268 A; 187 C; 213 G; 194 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1-7e-74 Length: 862
 Score: 759.50 Matches: 152
 Percent Similarity: 76.30% Conservative: 9
 Best Local Similarity: 72.04% Mismatches: 29
 Query Match: 66.74% Indels: 21
 Gaps: 5
 DB: 21
 US-10-021-811-36 (1-206) x AAC57336 (1-862)
 QY 1 MetaAspLysLys-----GlnGlnCysLysThrSerGlnAspProGluValArgLysGly 18
 DB 31 ATGGACAAAGAAAGCCAGACGACAGCATGTAATCCCAAGATGTCGAGGTGAGAAAAGG 90
 QY 19 ProTPrHrMetGluGluAspLeuIleuMetAsnTyrIleAlaAsnHisGlyGluGly 38
 DB 91 CCGTGGACGATGGAAGAGATCTATCTCATCAACTACATACAGCAATCAAGCGCGAAGGC 150
 QY 39 ValTrpAsnSerIleuAlaLysAlaGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
 DB 151 AGTTGGAACCTCCCAAGCAAGCTGCTGTTAAACGTACCGGGAAGAGTTCGGCTC 210
 QY 59 ArgTrpLeuAsnTyrLeuArgProAspValArgArgLysAsnIleHrProGluGluGln 78
 DB 211 CGGTGGCTGAACCTATCTGCGACCGACGTCGAGAGGCAACATCACTACGAGAGCAG 270
 QY 79 LeuLeuIleMetGluLeuHisAlaLysTrpLysAsnArgTrpSerLysIleAlaLysHis 98
 DB 271 CTCCTGATCATGGAATGTCATGCCAAGTGGGAAACAGTGTCTTAAATTCGCAAGCAT 330
 QY 99 LeuProGluArgTrpAspAsnGluIleLysAsnTyrTrpArgTrpArgGlnLysHis 118
 DB 331 CTCCTCGGAAGAGACTGCAATAGATTAAGAATCTTGGAGGACTTAACATCCAAACAC 390
 QY 119 IleLysGlnAlaGluAsnPhgGlnGlnInSerSerAsnSerGluIleAsnAspHis 138

DB 391 ATCAGCAAGCAGACAGGCTTCTCTGTCAGACGCTCCGAGATGAGTAT----- 438
 QY 139 GlnAlaSerThrSerHisValSerThMetAlaGluProMetGluMetLys---SerPro 157
 DB 439 CAAGCAAGCAGACCAAGCCATGATGTCAGACCCGATGAGACCTACGACTCCAGC 498
 QY 158 ProCysTrpGlnGly-----MetLeuGluProPheSerThrGlnPheProThrIleAsn 175
 DB 499 CCGTCATCTCCAGGCAACCAACATGAGACCTTTCGCGGTGAATTTCGCGTC----- 552
 QY 176 ProAspGlnSerSerCysCysThrAsnAspAsnAsnIleAsnTyrTrpSerMetGlu 195
 DB 553 -----GAGTCAAAATGAAGCCTACTACGAGCATGAGAC 582
 QY 196 AspSerTrpSerMetGluLeuAsnGlyAsp 206
 DB 583 GATCTTGTGCTATGAGGAGTACTCAATGGCGAT 615
 RESULT 2
 ID AAC40587 standard; DNA; 910 BP.
 AC AAC40587;
 XX
 DT 17-OCT-2000 (First entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 28847.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KM protein identification; signal transduction pathway;
 XX metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 XX
 PN EF1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0134941.
 PR 21-MAY-1999; 99US-0135124.
 PR 24-MAY-1999; 99US-0135353.
 PR 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142155.
PR 06-JUL-1999; 99US-0142035.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161008.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.


```

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144337.
PR 19-JUL-1999; 99US-0144338.
PR 19-JUL-1999; 99US-0144339.
PR 19-JUL-1999; 99US-0144339.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

```

```

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Score: 3,226-64 Length: 908
Percent Similarity: 667.00 Matches: 132
Best Local Similarity: 72.43% Conservative: 23
Query Match: 58.61% Mismatches: 51
DB: 21 Indels: 8
Gaps: 4

```

US-10-021-811-36 (1-206) x AAC48506 (1-908)

```

QY 1 MetAspLysLysGlnGlnCys-----LysThrSerGlnAspProGluValArgLysGly 18
DB 108 ATGAGAGAAAAGAGAAAGTGGTGGTCTGGATCGAGAGATGAGAGATGAGAAAGG 167
QY 19 ProTriPthMetGluGluAspLeuIleLeuMetAsnTYRlleAlaAsnHISglYglucly 38
DB 168 CCATGACGACATGAGAAAGATTTGATTCATCATTAATGCCAATCATGTCGAAGCT 227
QY 39 ValTrpAsnSerLeuAlaIysAlaIagIleuLysAlaGngLYLysSerCysArgLeu 58
DB 228 GTTGGAACTCTCCGCAAAATCTGCGAGACTAAACGACCGGAAAGTTCGGGCTC 287
QY 59 ArgTrpLeuAsnTYRLeuArgProAspValArgArgLYAsnIleThProGluGln 78
DB 288 CGGTGGCTGACTACCTCCGACCTGATGTGCGAGCGGAAATATACACCGAGAAACAG 347
QY 79 LeuLeuIleMetGluLeuHISAlaIalYtrpGlyAsnArgTrpSerLysIleAlaLysHIS 98
DB 348 CTCACCATCAWGGAGACTCATGCAAAATGCGAAATAGTGCTCAAAAATTCACAAACAT 407
QY 99 LeuProGlyArgThrAspAsnGluIleLysAsnTYRTrpArgTrpArgIleGlnLysHIS 118

```


KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PN
XX
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123348.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0128845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 14-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 18-MAY-1999; 990S-0134370.
PR 19-MAY-1999; 990S-0134768.
PR 20-MAY-1999; 990S-0134941.
PR 21-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139434.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139500.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 23-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.

PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142035.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144003.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 23-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.

Accession No.:	2_29e-55	Length:	621
Score:	585.00	Matches:	121
Percent Similarity:	65.14%	Conservative:	21
Best Local Similarity:	55.50%	Mismatches:	44
Query Match:	51.41%	Indels:	32
DB:	21	Gaps:	6

US-10-021-811-36 (1-206) x AAC41714 (1-621)

QY	1 Metasplyrlysglngnclyls-----thrsrlnaspprogu-----	14
Db	13 ATGAGAAAGAAAGGAGAGTAAAGCCAGCAATACGTACAGAAAGAAAGAAAGAACAA	72
QY	15 ValArgIySgIyPrIcRPrThmetIcIuIuSpIeuIleuMetAsnYrIleAlaAsn	34
Db	73 GTGAGAAAAGGACCTTGGACATATGGAACAAACATTTTCATCTTATTTACATCCTTAAT	132
QY	35 HisGlyIcIuIyValITrAsnSerIeuAlaIyAlaIcIyIeuIyAsrAGsngIyLys	54
Db	133 CATGGGAAGGCTTTGGAACTCTGTCCCAAGAGCTCTGGTCTAAACACTACGGAAA	192
QY	55 SerCysArgIeuAIrIPrIeuAsnYrIleuArIpprObAspValArgArgIyAsnIlethr	74
Db	193 AGTTGGCGGCTCGGAGCGGTAAATATCTCCAGACAGATGTGGCGGAGGAAACATAAAC	252

QY	75	ProgluIngluInleuIleuIleMeGluIleuHisIalysTrpGlyAsnArqTrpSerLys	94
Db	253	GCAGAGAAACAGCTTTTGATATTGATTCATCTTCAAGCTTGGAACAGAGTGTGCAG	312
QY	95	IleAlaLysHisIleuProGlyArgTrhAspAsnGluIleLysAsnTyTrpArgTrhArg	114
Db	313	ATTGGAGAGCATCTTCCGCGAAGAACGACACAGAGATAAGAAACTTCTGGAGACAAAG	372
QY	115	IleGluLysHisIleLysGlnAlaGlu-----AsnPheGlnGlnInLysSer	130
Db	373	ATTCCAGAGACACATCAAGAGTGTCTATCCGAAATAATGATGAATCTCAACATCATTTGCG	432
QY	131	AsnAsnSerGluIleAsn-----AspHisGlnAlaSerThrSerHisValSerThiMet	148
Db	433	GGAAACTCACAAGAGCTCCGGGATGCAGCAGCAGAGCTCCGCGCAAGCCATGACACAG	492
QY	149	AlaGluProMetGluMetLysSerProProCysTyGlnGlyMetLueGluInProPheSer	168
Db	493	GCTGAGAGCTTCTCT-----CAAGCGAAGACG	519
QY	169	ThrGlnPheProTrhIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsn	188
Db	520	ACGAGCTTTAATGTGTGCT-----GAACAAACACTCAAAAC	552
QY	189	IleAsnTyTrpSerMetGluAspSerTrpSerMetGluLeuLeuAsnGlyAsp	206
Db	553	GAGAAATTACTGAGACGTTGAGATCTGTGCCCGTCCACTTGCTTAATGCTGAC	606
RESULT 6			
AAC37953			
ID	AAC37953	standard; DNA; 1127 BP.	
XX			
AC	AAC37953;		
XX			
DT	17-OCT-2000	(first entry)	
XX			
De	Arabidopsis thaliana	DNA fragment SEQ ID NO: 19258.	
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
XX	metabolic pathway; promoter; termination sequence; ss.		
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	188
		99US-0123180.	
PR	05-MAR-1999;	99US-0122548.	
		99US-0125788.	
PR	23-MAR-1999;	99US-0126264.	
		99US-0126785.	
PR	29-MAR-1999;	99US-0127462.	
		99US-0128234.	
PR	06-APR-1999;	99US-0128714.	
		99US-0128845.	
PR	16-APR-1999;	99US-0130077.	
		99US-0130449.	
PR	21-APR-1999;	99US-0130510.	
		99US-0130891.	
PR	23-APR-1999;	99US-0131449.	
		99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
		99US-0132485.	
PR	04-MAY-1999;	99US-0132486.	
		99US-0132487.	
PR	06-MAY-1999;	99US-0132483.	
		99US-0132456.	
PR	11-MAY-1999;	99US-0132456.	

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

D	b	182	GAGCAGGCTTGTATTCCTCGACACTCCATTCCCGCTGGGGCAATGATGTCGAAGAATGCCG	241
O	y	97	LysHisLeuProGlyValArgThrAspAsnGluIleLysAsnTyrrTrpArgThrArgIleGln	116
D	b	242	CAAACACTTGCCGGGCAGGACCAGCAACAAGATCAAAGAACTACGTGGGAACCCGGTGCAG	301
O	y	117	LysHisIleLysGlnAlaGluAsnPheGlnGlnGlnInsertSerAsnAsnSerGluIleAsn	136
D	b	302	AAGCACGGCAGGACAGCTC-----AAATGTGACCTCAACAGCAAGCATTTCAAG	349
O	y	137	AspHis-----GlnAlaSerThrSer	143
D	b	350	GACGCCGATGAATACCTGTGATGCGAGGCTGTGTAGAGATCCAAGCGGCTCTCCGC	409
O	y	144	HISVALSerThr-----MetAlaGluProMetGluMetTyrrSer	156
D	b	410	TCTGTCTGCAGCCGCTACTGTGCCCCGCCGCCGATGGCAGACCCCACCACCAATGGCCACC	469
O	y	157	-----ProProCysTyrrGlnGlyMetLeuGlu	165
D	b	470	ACCCGACGATCCAACATTCGCGCGCATGGCTTTCGCCGCCGCCCTCGCGGCGATGGCGGC	529
O	y	166	ProPheSer-----ThrGlnPheProThrIleAsnProAspGlnSerSerCys	181
D	b	530	GACTTCAGGGCGGGCGAGTAGTAATGTGGCCGCCAGCTACAGACACCCGGGAACCTCCGC	589
O	y	182	CysThrAsnAspAsnAsnAsnIleAsnTyrrTrpSerMetGluAspSerTrpSerMetGln	201
D	b	550	ACGACG-----GCGCTTCCGACATCATTCGCTGGCGAG	622
O	y	202	Leu	202
D	b	623	CTC	625
RESULT 8				
AAD05784	ID	AAD05784	standard; cdNA; 1137 BP.	
XX	AC	AAD05784;		
XX	DT	31-JUL-2001	(first entry)	
XX	DE	Arabidopsis thaliana	transcription factor, G1324 cDNA.	
XX	KW	Transcription factor; biochemical characteristic; controlling element;		
KW	KM	structural characteristic; developmental characteristic; gene therapy		
KW	KM	agricultural biotechnology; plant trait modification; ss.		
XX	OS	Arabidopsis thaliana.		
XX	FH	Key	Location/Qualifiers	
FH	FT	CDS	54..914	
FT	FT	/tag= a		
FT	FT	/product= "transcription factor, G1324"		
PN	XX	M0200136597-A1.		
XX	PA	14-NOV-2000; 2000MO-USJ1344.		
XX	PD	25-MAY-2001.		
XX	PF	14-NOV-1999; 99US-0166228.		
XX	PR	17-APR-2000; 2000US-0197899.		
PR	PR	22-AUG-2000; 2000US-0227439.		
PA	PA	(MEND-) MENDEL BIOTECHNOLOGY INC.		
PA	PA	(CREE/) CREELMAN R.		
PA	PA	(YUG/) YU G.		
PA	PA	(ADAM/) ADAM L.		
PA	PA	(RIEC/) RIECHMANN J L.		
PA	PA	(HEAR/) HEARD J.		
PA	PA	(SAMA/) SAMARA R.		
PA	PA	(PIIG/) PILGRIM M.		

PA	(PINE)/PINEDA O.
PA	(JIANG)/JIANG C.
PI	Greelman R., Yu G., Adam L., Riethmann Jr., Heard J., Samaha R;
PI	Pilgrim M., Pineda O., Jiang C;
XX	
DR	WPI: 2001-33599/35.
DR	P-PsDB; AAE01900.
XX	
PT	Nucleic acids encoding plant transcription factor polypeptides, useful
PT	for altering the biochemical characteristics of plants e.g. corn,
PT	potato and cotton plants -
PS	Claim 4; Page 87-88; 127pp; English.
XX	
CC	The present sequence is Arabidopsis thaliana transcription factor,
CC	G1324 cDNA. The transcription factor is used for altering a plant's
CC	biochemical characteristics. The transcription factor may be used to
CC	alter the structure and developmental characteristics of plants such as
CC	soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC	alfalfa, sugar cane, turf, banana, blackberry, blueberry, strawberry,
CC	raspberry, cantaloupe, carrot, cauliflower, coffee, cucumber, eggplant,
CC	grapes, honey dew, lettuce, mango, melon, onion, papaya, peas, peppers,
CC	pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon,
CC	rosaceous fruits and/or vegetable brassicas. Transcription factors are
CC	key controlling elements of biological pathways and altering expres-
CC	sion levels of 1 or more transcription factors can change entire biolog-
CC	ical pathways in an organism. Therefore manipulating transcription fac-
CC	tors in plants offers great potential in agricultural biotechnology
CC	for modifying a plant's traits. Transcription factor cDNA is useful in
XX	gene therapy.
XX	
SQ	Sequence 1137 BP; 378 A; 213 C; 233 G; 313 T; 0 other;
Alignment Scores:	
Pred. No.:	6.63e-50 Length: 1137
Score:	539.00 Matches: 107
Percent Similarity:	66.18% Conservative: 28
Best Local Similarity:	52.45% Mismatches: 43
Query Match:	47.36% Indels: 26
DB:	22 Gaps: 5
US-10-021-811-36 (1-206) x MAD50784 (1-1137)	
OY	1 MetasplyslsgnglngncylsystrSerGlnaspProgluValArglysglyProtp 20
Db	::: :::: :::
66	ATGAAGAAGAAAGAACGTTCACAAAGAAAGTGAAGTAAGAAACAAGAGCGCCTTG 125
OY	21 Thrmctlnclunspleuleleleuetsnryrllleaahnsiglygluglyval 40
Db	::: :: ::: ::::
126	ACTTTGAGGAGAACACACACTCTTCACAAAATTCAATCCCTCAATACGGGTGCCTTG 185
OY	41 AsnserleualalysalaiaaglyleuysarGansgnylysserCyatgleuAvtgp 60
Db	::: :: ::: :::
186	AATACGTGCCCAAATGTGCTGGCTTAAGAGAACTGGGAAAAATTTAATTGAGATGG 245
OY	61 leuasntyrleuarGproaspvalArGARglyasnllerhrProglungluInleu 80
Db	::: :: ::: :::
246	TTCGAATTACTTGAACCACGACATAAGACGAGGAATCTTACTCTCAAGAACAGCTTTTG 305
OY	81 lleueticluenuhislaIalysrtrpglyasnArGrprpserylleAlushileupro 100
Db	::: :: ::: :::
306	ATTCCTTAGCTTACACTCTAAATGGGTAATAGTGTCTCAAGATGACAGTACTTGCCA 365
OY	101 GlyargThrAspaNgiluilelysantYrrpArghThrArglllegluNshisley 120
Db	::: :: ::: :::
366	GGAAGAACGATTAACAGATCAAGAACTATTGGAGAACAGAGTCCAAAAACAAGCTGT 425
OY	121 GlhalaglunsnheglncInlserserissnsersguilleksnasphsiglna 140
Db	::: :: ::: :::
426	CMACTC---AACATCGAA-----TCTAACAGCGCACATTCCTTTTGAC--GCTGTP 470
OY	141 SerThserHisValserThrmelaagluPrometglumetyrSerProprocystyr 160

Db 471 CGTAGTTTGGTCCCTACATGATGAGAGATGAGAAACATCA----- 518
Qy 161 GlnGlyMetLeuGluProPheSerThrGlnPheProThrIleAsnProAspGlnSer 180
Db 519 -----TCCACTACTACTACTTAT 536
Qy 181 CysCysThrAsnAspAsnAsnIleAsnTyrTrp-----SerMetGluAspSer 197
Db 537 TGTGTGTCGCCAAACACACACACACTCTCTCTCTCTCTCAATCTCACGACTCT 596
Qy 198 TrpSerMetGln 201
Db 597 TTAAGTATGCAA 608
RESULT 9
ABK65252
ID ABK65252 standard; cDNA; 981 BP.
XX
XX ABK65252;
AC
XX 02-JUL-2002 (first entry)
XX
XX Arabidopsis cDNA encoding a transcription factor #104.
XX
XX Plant; ss; gene; transcription factor; transgenic;
KW agriculture; metabolic chemical; environmental stress; drought;
KW microbial disease resistance; herbicide resistance; seed yield;
KW fruit yield; growth rate; leaf senescence; flower senescence.
XX
XX Arabidopsis thaliana.
XX
XX W0200215675-A1.
XX
XX 28-FEB-2002.
XX
XX 22-AUG-2001; 2001WO-US26189.
XX
XX 22-AUG-2000; 2000US-227439P.
XX
XX 16-NOV-2000; 2000US-0713994.
XX
XX 16-APR-2001; 2001US-0837944.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX (PIVJ/) PILGRIM M.
XX
XX (CREE/) CREELMAN R.
XX
XX (DUBE/) DOBELLA A J.
XX
XX (HEAR/) HEARD J.
XX
XX (JIAN/) JIANG C.
XX
XX (KEDD/) KEDDIE J.
XX
XX (ADAM/) ADAM L.
XX
XX (RATC/) RATCLIFF O.
XX
XX (REUB/) REUBER J L.
XX
XX (RIE2/) RIECHMANN J L.
XX
XX (YUG5/) YU G.
XX
XX (PINE/) PINEDA O.
XX
XX Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J;
XX
XX Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
XX
XX WPI: 2002-293022/33.
XX
XX P-PSDB; AA093066.
XX
XX An isolated or recombinant polynucleotide used to produce a transgenic
XX
XX plant -
XX
XX Claim 4; Page 443-445; 941pp; English.
XX
XX The invention relates to 1 of 232 isolated or recombinant polynucleotides
XX
XX encoding an Arabidopsis thaliana transcription factor, their variants,
XX
XX complements, fragments, or related polynucleotide with 31% to 95%
XX
XX sequence identity, where the plant possesses an altered trait as compared
XX
XX to a wild-type or reference plant, or the plant exhibits an altered
XX
XX phenotype as compared to a wild-type or reference plant, or the plant

CC exhibits ectopic expression or altered expression of one or more genes
CC associated with a plant trait as compared to a wild plant. Also included
CC are a transgenic plant comprising the polynucleotides, a computer
CC readable medium having stored sequence information, and identifying a
CC homologue sequence from a database comprising a plurality of known plant
CC sequences comprising inputting sequence information selected from one of
CC 464 fully defined sequences given in the specification. The isolated or
CC recombinant polynucleotide is used for producing a plant having a
CC modified trait, the method comprising selecting a polynucleotide that
CC encodes a polypeptide or an antisense nucleic acid, inserting the
CC polynucleotide or antisense nucleic acid into an expression vector,
CC introducing the vector into a plant or a cell of a plant to overexpress
CC the polypeptide or antisense nucleic acid, thereby producing a modified
CC plant, and selecting for a modified trait (e.g. increased
CC production of agriculturally useful proteins or metabolic chemicals,
CC pest tolerance, environmental stress response (e.g. drought), microbial
CC disease resistance, herbicide resistance, seed and fruit yield, growth
CC rate, leaf and flower senescence and many other traits listed in the
CC specification). The present sequence is one of the 232 polynucleotides
CC encoding an A. thaliana transcription factor.
XX
XX SQ Sequence 981 BP; 319 A; 193 C; 235 G; 234 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2,81e-48 Length: 981
XX Score: 523.50 Matches: 119
XX Percent Similarity: 55.77% Conservative: 26
XX Best Local Similarity: 45.77% Mismatches: 32
XX Query Match: 46.00% Gaps: 83
XX DB: 24 Indels: 9
XX
XX US-10-021-811-36 (1-206) x ABK65252 (1-981)
Qy 8 LysThrSerGlnAspProGluValArgLysGlyProTyrPheMetGlnLysPheLeu 27
Db 90 AAGGTGAGAGAAAGAAATGAGCGCTAAGAGACAGCTCCATGAGAGATTGATGAGC 149
Qy 28 LeuMetAsnTyrTlleAlaAsnHisGlyGluGlyValTyrPheSerLeuAlaLysAla 47
Db 150 CTCATCAATTACATTCCTGTCATGCTGAGAGGTCGATGAGATCTCTCTGCTGGCGC 209
Qy 48 GlyLeuLysArgAsnGlyLysSerCysArgLeuArgTyrPheAsnTyrLeuArgProAsp 67
Db 210 GAACCTCAAAAGAGACCGGAAAGAAACCTGCACACTTCGGTGCATGATCTCGACACAGAT 269
Qy 68 ValArgArgGlyLysAsnIleThrProGluGluGlnLeuLeuIleMetGlnLeuHisAlaLys 87
Db 270 GTGCGCGGTGAGAAACATTAACCTCGAAGAACAACTCTGATTCCTGAACCTTCACACACT 329
Qy 88 TrpGlyAsnArgTyrPheSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIle 107
Db 330 TGGGGCAATGATGCTTAAGATTGACCAATATTTCAGGAAGAAAGGATTAACGAGATC 389
Qy 108 LysAsnTyrTyrPheArgThrArgIleGlnLysHisIleLysGlnAlaGlu-----Asn 124
Db 390 AAAAATGATTGGAGAACACACTGTTCAAAAGACATGCAAAACAGCTTAAGCGACGTGAAC 449
Qy 125 PheGlnGln----- 127
Db 450 AGTCACAAATTTAAAGACACCATGAGATCTTTGGATGCTCGCTAGAAAGATC 509
Qy 128 -----GlnSerSerAsn 132
Db 510 CAAAGCCCGTCATCGGGTGTGTCATGTCATCTTGGGTGACCAACCTCCATGATCAG 569
Qy 133 SerGluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGluProMet 152
Db 570 TTGCTGATCAACAACAAC-----AACACCAACAAGCTGATATTTGGCT----- 614
Qy 153 GluMetLysSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPhePro 172
Db 615 TTAATGAGTAACCTTATGCTTAC----- 638

woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, cedar, manogany species or to modify the activity of a polypeptide in a plant

CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and MYB.

XX Sequence 524 BP; 143 A; 131 C; 137 G; 113 T; 0 other:

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1.29e-43	524	478.00	81	20	18	0	0
Percent Similarity:	84.87%						
Best Local Similarity:	68.07%						
Query Match:	42.00%						
DB:	21						

US-10-021-811-36 (1-206) x AAC57194 (1-524)

OY 3 LysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrpThrMet 22
 DB 154 AGGAAATGAGCATGTCGCCGAGAGAGAGGCTGACCTGCCAAGGGGCGCATGACCTGC 213
 OY 23 GlnGluAspLeuIleLeuMetAsnTyrTlleAlaAsnHisGlyGluGlyValTrpAsnSer 42
 DB 214 GAGGAGACAAATTTGCTCATCTCATCGATCCACATCCACGGGAGGAGCGCTGGAAATG 273
 OY 43 LeuAlaLysAlaLagLysLeuLysAsnTyrLysSerCysArgLeuArgTrpLeuAsn 62
 DB 274 TTGGCCAGAGCGCAGATGTAAGAGAACTGCAAAAGCTGACAGATTAAAGTGGCTGAT 333
 OY 63 TyrLeuArgProAspValArgArgLysAsnIleThrProGluGlnGluLeuIleMet 82
 DB 334 TACCTGAGACCCGACATCAAGCGCGGAAATCTCACCCGCAAGAACACTCATGATCTCT 393
 OY 83 GlnLeuHisAlaLysTrpGlnAsnArgTrpSerLysIleAlaLysHisLeuProGlyArg 102
 DB 394 GAATCTCACCACAAATGGGGCAACAGTGTGTTGAAAATCGCCAGTATCTCCAGAGAG 453
 OY 103 ThrAspAsnGluIleLysAsnTyrTrpArgTrpArgIleGlnLysHisIleLysGln 121
 DB 454 ACAGATACAGAGATCAAGAACTACTGAGAGCGGCGTGCAGAAACAGCGCGCAG 510

RESULT 13

AAX25572 standard; cDNA; 837 BP.

XX AAX25572;

XX 02-AUG-1999 (first entry)

XX Arabidopsis thaliana MYB2 (AtMYB2) cDNA.

KW MYB2 gene; AtMYB2 gene; transcription factor; transgenic plant;

KW environmental stress; stress tolerance; anaerobic stress;

KW drought; flooding; salt; cold; crop protection; ss..

XX Arabidopsis thaliana.

XX Key location/Qualifiers

XX FT CDS 10.831 /tag- a

XX MO916878-A1.

XX 08-APR-1999.

XX 25-SEP-1998; 98MO-AD00812.

XX 26-SEP-1997; 97AU-0009479.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

PI Dennis ES, Dolferus RAM, Hoeren FU, Peacock WJ;

XX WPI: 1999-263695/22.

DR P-PSDB; AAY05831.

XX Altering expression of a stress-related gene in a plant

XX Claim 22; Page 72-74; 11pp; English.

XX This is the sequence of Arabidopsis thaliana ecotype C24 cDNA
 CC encoding the stress-related protein AtMYB2 (see AAY05831). The
 CC cDNA was obtained by RT-PCR of anaerobically induced root RNA
 CC derived from ecotype C24. The inventors have demonstrated that a
 CC family of transcription factors, the MYB2 family of proteins, are
 CC capable of inducing the expression of stress-related genes in plant
 CC cells in response to stress situations. The MYB2 protein encoded
 CC by the AtMYB2 gene is useful for the purposes of inducing or
 CC repressing the expression of plant-expressible genes that are
 CC involved in the plant's response to anaerobic stress, flooding
 CC stress, cold stress, dehydration stress, drought stress, heat
 CC stress or salinity. The stress-related gene, such as the MYB2
 CC alcohol dehydrogenase Adh1 gene, is transactivated by the MYB2
 CC polypeptide by virtue of the presence of at least one copy of a
 CC cis-acting regulatory sequence, in particular a MBS-1 (Myb binding
 CC site) motif and/or a GC-motif and/or a G-box-1 motif and/or a
 CC G-box-2 motif, as found in the Adh1 promoter (see AAX25573-74). The
 CC MYB2 protein is able to target several stress-induced enzymes,
 CC rather than targeting just one, allowing the production of stress
 CC tolerant transgenic plants.

XX Sequence 837 BP; 246 A; 176 C; 212 G; 203 T; 0 other:

Alignment Scores:	Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.31e-43	837	475.00	92	27	36	24	4	0
Percent Similarity:	66.48%							
Best Local Similarity:	51.40%							
Query Match:	41.74%							
DB:	20							

US-10-021-811-36 (1-206) x AAX25572 (1-837)

OY 11 GlnAspProGluValArgLysGlyProTrpThrMetGlnGluAspLeuIleLeuMetAsn 30
 DB 55 GAAGATTCGATGATACGGAAGAGTCCATGAGACCGAGAGAGATGCAATCTGATCAAC 114
 OY 31 TyrTlleAlaAsnHisGlyGlnGlyValTrpAsnSerLeuAlaLysAlaLysGlyLeuLys 50
 DB 115 TTGCTCTATTCATGATGCGATGCTGCTGAGAACACATCGCTGCTGCTGCTGCTGCTG 174
 OY 51 ArgAsnGlyLysSerCysArgLysArgTrpLeuAsnTyrTlleAsnArgProAspValArg 70
 DB 175 CGAAGTGGTAGAGTGTGATGATTAGATGCTTATTTCTTACGTCGACGATTTAGAGA 234
 OY 71 GlyAsnIleThrProGluGlnGluLeuIleMetGlnLeuHisAlaLysTrpGlyAsn 90
 DB 235 GGCACATCACTCTCGAACAACATTTATGATCTCAAACTCCATCTCTTGGGCAAT 294
 OY 91 ArgTrpSerLysIleAlaLysHisLeuProGlyArgTrpAsnGlnIleLysAsnTyr 110
 DB 295 AGGTGTCGAGAGTTTCCGCAATTTCTACCGGAGAAACAGATTAATGAATTAATAT 354
 OY 111 TrpArgTrpArgIleGlnLysHisIleLysGlnAla-----GluAsnPhe 125
 DB 355 TGGAGAACTCGAGTCCAAAAGCAAGCAACACTAAGATCGATGTTAAAGTAATCTT 414
 OY 126 GlnGlnGlnSerSerAsnAsn-----SerGluIleAsnAspHis 138
 DB 415 TTCAGAGAGACTTGGAAATGTTTGGATGCCAGATTAGTGAACGATTAAGCCCAA 474
 OY 139 GlnAlaSerThr-----SerHisValSerThrMetAlaGluProMet 152
 DB 475 TCATTAACCCACAGTGTGAACAAGTGAAGTCAATGATGATCAGCAACCAAGTCAACAGTT 534

OY 153 j3UMeTyrSerProFroCysTyrGlnGlyMetLeuGluPProPhSerThrGlnPhe 171
Db 535 AACGACCGAGTCGCG-----GTCGACCGCGGTTTCGTCACATTC 573
RESULT 14
AAC44894
ID AAC44894 standard; DNA; 959 BP.
XX AAC44894;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44531.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD
XX 06-SEP-2000.
PF
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137252.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 18-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

```

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Alignment Scores:

```

Pred. No.: 3.8e-41
Score: 459.00
Percent Similarity: 54.22%
Best Local Similarity: 41.78%
Query Match: 40.33%
DB: 21

```

```

Length: 959
Matches: 94
Conservative: 28
Mismatch: 57
Indels: 46
Gaps: 5

```

US-10-021-811-36 (1-206) x AAC44894 (1-959)

OY 16 Arglysllyprotrprrhmetglugluaspheulleuemetasnyrrilealaasnhis 35

```

DB 204 AGAAAGGCGCTTGACCTGCTGTAAGAGACCGCGCTTTGATGCTATGTCACCTTCAC 263
OY 36 GLYGLUGLYVALTRPAsnSerleuAlaIysAlaIglYleuYAsnGlyLysSer 55
DB 264 GGTGAAGGCGCATGGAACCTCTGTCGAGCGCCCGGGGCTTGGAAGAAATGGGAAGAC 323
OY 56 CysArgLeuAlaGTrpLeuAsnTrpYleuArgProAspValAlaGArgGlyAsnIleThrPro 75
DB 324 TGCAGGTAAAGTGGGTAACTACTGTAAGACCGACCTCAAGAGAGCAATATCTCT 383
OY 76 GLUGLUGLLeuLeuIleuMetCLeuIleuHisAlaYsTrpGlyAsnArgTrpSerIle 95
DB 384 CATGAAGAAGAACCATATCTCTGCTTACCTGCTAGTGGCGCAATAGGTGCTCCACATT 443
OY 444 GCACGTAGTTTACCGGGAAGACAGCAACAAATCAAGAACTATTGCGCAACCATTTTC 503
DB 116 GlnYsHisIleYsGlnAlaGluAsn----- 124
OY 504 AAGAAAGACGAGAGTCTCCACTTAACAGTGGCGGAAGAACAAAGAACCAATCTTGAG 563
DB 125 -----PheGlnGlnGln----- 128
OY 564 AGCAACAATTTCCAGCAGCAAGACAAATGAGTTGCAGCAAGAACACACTTCTTCAA 623
DB 129 -----SerSerAsnAsnSer 133
OY 624 TTCATCAATTCGACATGAAAAAGATCATGTCCTTACTAGTACGACACAAACAAATGCT 683
OY 134 GluIleAsnAspHisGlnAlaSerThrSerHisValSerThrMetAlaGluProMetGlu 153
DB 684 GATTAACACCTTCAGCAGTAGCAGTAGTGGCAATGAGCATTTATGTATCTACCATCAG 743
OY 154 MetTyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThr---GlnPhePro 172
DB 744 ATCAACATTCACAAACAACTCTGCT---TGTGAACCAATATGATACGGGTATTACCG 800
OY 173 ThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsnAsnIle-----Asn 190
DB 801 GTGTTCCGGTAAACATTAACCTGAGCTATGTAATGAATGAATGAATGATTTGGAGCGT 860
OY 191 TyrTrpSerMetCLeu 195
DB 861 TTATGGAATCTGGAT 875

```

RESULT 15
AAC56152
AAC56152 standard; DNA; 389 BP.

AAC56152:

25-JAN-2001 (first entry)

Eucalyptus grandis transcription factor DNA sequence #283.

KW plant: transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor; MADS;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox;
KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.

Eucalyptus grandis.

WO200053724-A2.

14-SEP-2000.

09-MAR-2000; 2000WO-US06112.

PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLEET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 XX Wood M, McGrath A, Shenk MA, Glenn M;
 PI
 XX WPI: 2000-579369/54.
 DR

XX New isolated polynucleotide encoding a plant transcription factor for
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
 PT having modified gene expression or modified activity of a polypeptide
 PT

XX
 PS Claim 1: Page 120; 747pp; English.

XX The present invention relates to novel plant transcription factors from
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
 CC sequence for one such transcription factor. The transcription factor may
 CC be used to produce a plant having modified gene expression such as a
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
 CC mahogany species or to modify the activity of a polypeptide in a plant.
 CC The transcription factors of the present invention are members from the
 CC following families of regulatory proteins: bZIP, bZIP family of G-box
 CC binding factors, basic helix-loop-helix zipper, LIM domain, AP2
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
 CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
 CC and MYB.

XX
 SQ Sequence 389 BP; 112 A; 102 C; 97 G; 78 T; 0 other;

Alignment Scores:

Pred. No.:	1,75e-37	Length:	389
Score:	421.00	Matches:	81
Percent Similarity:	88.17%	Conservative:	1
Best Local Similarity:	87.10%	Mismatches:	9
Query Match:	36.99%	Indels:	2
DB:	21	Gaps:	1

US-10-021-811-36 (1-206) x AAC56152 (1-389)

```

QY 1 MetAspLysLys-----GInGInCysLysThrSerGInAspProGluValArgLysGly 18
   |||||
Db 44 ATGGACAAGAAAGCCGACGACGACAGTGTAGTCCCAAGATGTCGAGGTGAGAAAGGG 103
   |||||

QY 19 ProTrpThrMetGluGluAspLeuIleLeuMetAsnTyrIleAlaAsnHisGlyGluGly 38
   |||||
Db 104 CCGTGGACGATGAGACGATCTCATCTCATCACTACATAGGATCAAGGGAAGGC 163
   |||||

QY 39 ValTrpAsnSerLeuAlaLysAlaLysGlyLeuLysArgAsnGlyLysSerCysArgLeu 58
   |||||
Db 164 AGTTGGAACTCCCTAGCCCAAGCTGCTGCTAAACGTACCGGAGAGAGTGTCTCGGCTC 223
   |||||

QY 59 ArgTrpLeuAsnTyrLeuArgProAspValArgArgGlyAsnIleThrProGluGluGln 78
   |||||
Db 224 CGGTGGCTGAACTATCTGCGACCGGAGTCCGGAGAGCAACATCACTACTAGAGGACAG 283
   |||||

QY 79 LeuLeuIleMetGluLeuHisAlaLysTrpGlyAsnArg 91
   |||||
Db 284 CTCTGATCATGGAATGCAATGCCAAGTGGGAAACAGG 322
   |||||

```

Search completed: February 18, 2003, 10:37:54
 Job time : 223 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 18, 2003, 10:31:30 ; Search time 1688 Seconds
(without alignments)
1976.463 Million cell updates/sec

Title: US-10-021-811-36
Perfect score: 1138
Sequence: 1 MDKKQCKTSDPEVRKGPV.....NNINYSMEDSMQMLNGD 206

Scoring table:
BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+.p2n.model -DEV-xlh
-O/cgn2_1/USPTO.spool/US10021811/runat_12022003_090133_9566/app_query.fasta_1.391
-DB-EST -QPMT-fastlap -SUFFIX-p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=100 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10021811.ecgn_1_1_899.ctunal_12022003_090133_9566 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NCG_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-YGAPOP=10 -YGAPEXT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estda:*
2: em_estlum:*
3: em_estln:*
4: em_estlu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_liv:*
20: em_gss_pln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	991	87.1	562	14	BM732121	BM732121 sal13b11..
2	859.5	75.5	635	14	BM046360	BM046360 NF047E06F
3	851	74.8	501	13	BM5227506	BM5227506 sal163906
4	846	74.3	782	10	BE658316	BE658316 GM700005B
5	816	71.7	552	13	BM527774	BM527774 sal165903
6	813.5	71.5	560	14	BM732539	BM732539 sal165908
7	808	71.0	669	12	BM457971	BM457971 NF037A10P
8	796	69.9	682	14	BM046831	BM046831 NF028G06F
9	795.5	69.9	568	13	BM522764	BM522764 sal164d10
10	784	68.9	523	9	A1930997	A1930997 sb45h07.y
11	779.5	68.5	544	13	BM527508	BM527508 sal162f08
12	774.5	68.1	542	13	BM528383	BM528383 sal157f09
13	774.5	68.1	542	13	BM528383	BM528383 sal157f09
14	724	63.6	681	9	A1486576	A1486576 EST244897
15	709.5	62.3	476	10	BE057370	BE057370 sn02c04.y
16	673.5	59.2	592	10	AW928296	AW928296 EST307029
17	671	59.0	558	10	BE324639	BE324639 NF024C04P
18	670.5	58.9	470	10	AW156542	AW156542 se28d11.y
19	657	57.7	409	12	BE804790	BE804790 s445f11.y
20	654.5	57.5	425	10	AW423958	AW423958 sh58e05.y
21	642	56.4	503	14	BM044458	BM044458 g91307.e
22	639	56.2	500	9	A1897784	A1897784 EST267227
23	639	56.2	519	9	A1897784	A1897784 EST267227
24	637.5	55.0	502	14	BM046505	BM046505 f00568.e
25	627.5	55.1	447	13	BI273011	BI273011 NF097E02F
26	621	54.6	395	12	BE805071	BE805071 s837a07.y
27	614	54.0	450	10	BE058847	BE058847 sn33b01.y
28	601	52.8	431	10	AW459279	AW459279 sh22h08.y
29	595	52.3	586	12	BE820766	BE820766 GM700012A
30	593	52.1	715	10	BE658737	BE658737 GM700007A
31	529.5	46.5	542	9	AU238946	AU238946 AU238946
32	527	46.3	397	10	AW432364	AW432364 sh13a04.y
33	526.5	46.3	727	12	BE607379	BE607379 WHE2471_H
34	522.5	45.9	676	12	BM013207	BM013207 QG34A09.y
35	514	45.2	588	14	BE649523	BE649523 NF080G02E
36	513.5	45.1	766	10	BE659054	BE659054 GM700008A
37	513.5	45.1	635	10	AW032656	AW032656 EST276215
38	513.5	45.1	690	14	BM0147546	BM0147546 NF041H09F
39	511	44.9	605	13	BM233398	BM233398 BU233398
40	510	44.8	370	12	BE642441	BE642441 EST355917
41	510	44.8	371	9	A1487923	A1487923 EST246245
42	508	44.6	502	10	BE459188	BE459188 EST414480
43	506.5	44.5	566	13	BI424418	BI424418 s4f3a05
44	506	44.5	510	12	BP325282	BP325282 su20e03.y
45	501	44.0	536	13	BI701275	BI701275 sag56g08

ALIGNMENTS

RESULT 1
LOCUS BM732121
DEFINITION sal13b11.y1 Gm-c1061 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1061-4678 5' similar to TR:Q39028 Q39028 ATMYB2. ;, mRNA
ACCESSION BM732121
VERSION BM732121
KEYWORDS
SOURCE ORGANISM
EST.
soybean.
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE	1 (bases 1 to 562)
AUTHORS	Shoemaker, R., Keim,

BASE COUNT 229 a 121 c 117 g 168 t
 ALIGNMENT SCORES:
 Pred. No.: 9,866-91 Length: 635
 Score: 859.50 Matches: 167
 Percent Similarity: 85.29% Conservative: 7
 Best Local Similarity: 81.86% Mismatches: 18
 Query Match: 75.53% Indels: 12
 DB: 14 Gaps: 5
 US-10-021-811-36 (1-206) x B0146360 (1-635)

QY 1 MetaspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
 DB ATGGATTAATAA---CCATGCAACCTCATCTCAAGATCCTGAAGTGAAGAAAGGCGCATGG 93
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTrpIleAlaAsnHisGlyGluGlyValTrrp 40
 DB ACCATGGAAGAAAGCTGGATTATTAATCAATTTATTTGCAAAATCATGCTGAAGCTGTTGG 153
 QY 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
 DB AATTCCTTGAAGCAAGAGCTGCTGCTTAAACGTACAGAGAAAGTTGCGAGCTTCGATGG 213
 QY 61 LeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGluLeu 80
 DB TTTAACTACTTCTGCTCAATGTTAGAGAGGAAATTTACCTGAGAGAAACACTTTGG 273
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
 DB ATCATGGAACCTTCATCAAGAGGCGGAAATAGCTGCTCAAAATTCGCAAGAGCATCTTCCA 333
 QY 101 GLeuArgThrAspAsnGluIleLysAsnTrpArgTrrpArgIleGlnLysHisIleLys 120
 DB GGAAGACACGACATAGATTAAAGATTTTGGAGAGACATAGATACAAAGACACATTTAG 393
 QY 121 GlnAlaGlu-----AsnPhenGlnGlnSerSerAsnAsnSerGluIle 135
 DB CAAGTGTATCATCTCACCACCAATATTTTCCAAACAAATGCT-----TTAGAGATA 444
 QY 136 AsnAsp---HisGlnAlaSerThrSer-HisValSerThrMetAlaGluProMetGluMe 154
 DB AATGATCATCATCAATCAAGTATGATAGCCCAAGTTTCCACATTTGAGAACCAATGCAATC 504
 QY 154 TyrSerProProCysTyrGlnGlyMetLeuGluProPheSerThrGlnPheProThrIi 174
 DB TTACTCTCCACCTTCATPACCAAGAACTTTGGAGCATTTCCAGCTCAATTTCCAAACAT 564
 QY 174 eAsnProAspGln---SerSerCysCysThrAsnAspAsnAsnIleAsnTrpTrrpSe 193
 DB CATGTGATCATCATCTCAGTCAAGTTGTTGACAAATGACAAATCAACAAATTTATTTGG 624
 QY 193 TrMetGluAsp 196
 DB 625 CATGAGAGAT 634

RESULT 3
 LOCUS BM527606 501 bp mRNA linear EST 19-FEB-2002
 DEFINITION sal63g06.y1 Gm-cl061 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl061-3780 5' similar to TR:Q39028 Q39028 ATMYB2. ; mRNA
 sequence.
 ACCESSION BM527606
 VERSION BM527606.1 GI:18733434
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 501)

AUTHORS
 Shoemaker R., Keim P., Vodkin L., Espeland J., Corvett V., Khanna
 A., Bolla B., Maria M., Hillier L., Kucaba T., Martin J., Beck C.,
 Wylie T., Underwood K., Slepoe M., Theising B., Allen M., Bowers
 Y., Person B., Swaller T., Gibbons M., Page D., Harvey N., Schuck
 R., Riller E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann
 R., Waterston R. and Wilson R.
 TITLE
 JOURNAL
 COMMENT
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R./Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: c@resgen.com web site:
 www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 421.
 Location/Qualifiers
 1..501
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl061-3780"
 /clone_1db="Gm-cl061"
 /tissue_type="mature flowers of field grown plants"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+, Site_1: EcoRI, Site_2:
 XhoI; The cDNA library was constructed from mRNA isolated
 from mature flowers of field grown plants for the cultivar
 Raiden. Complementary DNA was synthesized from mRNA using
 a primer consisting of a poly(dT) sequence with a XhoI
 restriction site. EcoRI adapters were ligated to the
 blunt-ended cDNA fragments followed by XhoI digestion. The
 cDNA fragments were directionally cloned into the
 EcoRI-XhoI restriction site of the pBluescript vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (GibcoBRL). This library was constructed in the
 laboratory of Dr. Randy Shoemaker."

BASE COUNT 170 a 104 c 114 g 113 t
 ALIGNMENT SCORES:
 Pred. No.: 6,896-90 Length: 501
 Score: 851.00 Matches: 157
 Percent Similarity: 96.34% Conservative: 1
 Best Local Similarity: 95.73% Mismatches: 5
 Query Match: 74.78% Indels: 0
 DB: 13 Gaps: 0
 US-10-021-811-36 (1-206) x BM527606 (1-501)

QY 1 MetaspLysLysGlnGlnCysLysThrSerGlnAspProGluValArgLysGlyProTrrp 20
 DB ATGGATTAATAA---CCATGCAACCTCATCTCAAGATCCTGAAGTGAAGAAAGGCGCATGG 69
 QY 21 ThrMetGluGluAspLeuIleLeuMetAsnTrpIleAlaAsnHisGlyGluGlyValTrrp 40
 DB ACCATGGAAGAAAGCTGGATTATTAATCAATTTATTTGCAAAATCATGCTGAAGCTGTTGG 129
 QY 41 AsnSerLeuAlaLysAlaIleGlyLeuLysArgAsnGlyLysSerCysArgLeuArgTrrp 60
 DB AATTCCTTGAAGCAAGAGCTGCTGCTTAAACGTACAGAGAAAGTTGCGAGCTTCGATGG 189
 QY 61 LeuAsnTrpLeuArgProAspValArgArgGlyAsnIleThrProGluGlnGluLeu 80
 DB TTTAACTACTTCTGCTCAATGTTAGAGAGGAAATTTACCTGAGAGAAACACTTTGG 249
 QY 81 IleMetGluLeuHisAlaLysTrpGlyAsnArgTrrpSerLysIleAlaLysHisLeuPro 100
 DB ATCATGGAACCTTCATCAAGAGGCGGAAATAGCTGCTCAAAATTCGCAAGAGCATCTTCCA 309

OY 101 GLYArGThAspAnGluIleLysAsnTYTPArGThArgIleGlnIlyshIleLys 120
 DB 310 GGTAGGACATATATGATGACCAAGAACTATTTGGAGACGAGTCCAGAACCATCAAC 369
 OY 121 G1A1AG1uAnSpHegInGlnInSerSerAsnSerc1u1leAsnAspHISg1nAla 140
 DB 370 CAAGCTGAGAACTTTCAGCAAAATTAATTAACCTGAGATTAATGATGATCAACCAAGCT 429
 OY 141 SerTherIstValSerTherMetAlaGluProMetGluMetYsrSerProProCysTYR 160
 DB 430 AGCACTGACGATGTTTACCATGCTGAGAACCATGAGACCTATTCCTCCACCTTTTAT 489
 OY 161 GlnGlyMetLeu 164
 DB 490 CAAGGATGTTA 501
 RESULT 4
 BE658316/c 782 bp mRNA linear EST 24-MAY-2001
 LOCUS GM700005B10E4 Gm-r1070 glycine max cDNA clone Gm-r1070-1759 3'
 DEFINITION mRNA sequence.
 ACCESSION BE658316
 VERSION BE658316.1 GI:9984208
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 782)
 Vodka, L., Kelm, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Expanding, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.
 A Functional Genomics program for Soybean (NSF 9872565)
 Unpublished (1999)
 Other_ESTS: A1930997 corresponding to Gm-cl015-278 (5')
 Contact: Vodka, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois
 Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582
 Email: l-vodka@uiuc.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com
 Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers
 1..782
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="Gm-r1070-1759"
 /clone_1b="Gm-r1070"
 /note="The library Gm-r1070 is a sequence-driven, rerecked
 set of 9,216 clones selected from cDNA libraries from
 various tissues and stages of development of soybean that
 represent 2,639 sequences from immature cotyledons, 1,770
 from immature seed coats, 3,938 from flowers, and 869
 from young pods. The 5' ESTs of the source clones from
 the different libraries was used to select singletons, or
 a representative of each contig, which were rerecked to
 form library Gm-r1070. The cDNA clones of the rerecked
 Gm-r1070 library were then sequenced at the 3' end. The
 contig analysis to select unique genes was performed by
 the laboratory of Ernest Retzel, Center for Computational
 Genomics and Bioinformatics, University of Minnesota,
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
 Rerecking was performed by Genome Systems, St. Louis,

http://www.genomesystems.com, and 3' sequencing by the
 Keck Center for Comparative and Functional Genomics,
 University of Illinois,
 http://www.life.uiuc.edu/bio-tech/Keck.html. Note: The
 corresponding 5' EST from each clone in the Gm-r1070
 library is listed in the 'OTHER EST' field. The detailed
 information on the source library for each clone can also
 be obtained by referring to the Genome Systems clone ID of
 the original cDNA library that is also listed under
 'OTHER EST'.
 BASE COUNT 215 a 144 c 143 g 253 t 27 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.25e-89 Length: 782
 Score: 846.00 Matches: 162
 Percent Similarity: 90.61% Conservative: 2
 Best Local Similarity: 89.50% Mismatches: 14
 Query Match: 74.34% Indels: 3
 DB: 10 Gaps: 2
 US-10-021-811-36 (1-206) x BE658316 (1-782)
 OY 29 MetAsnTYRlIeAlaAsnHISglYglYAlTYrPaSnSerLeuAlaAlaGly 48
 DB 780 ATCACTATATTCGCAAAATCATCTNNNNNGGTGTGANNNNCTGGCCAAAGCTGCTMNT 721
 OY 49 LeuLysArgAsnGluYsrSerCysArgLeuArgTYrPleAsnTYrLeuArgProAspVal 68
 DB 720 CTCAAACGATACCGGAAAGAGTTCGCCGCTTAAGTGGCTTAACCTCCGCTCATATTT 661
 OY 69 ArgArgGlyAsnIleThrProGluGluGluLeuIleMetGluLeuHisAlaLysTrp 88
 DB 660 AGAAGGGGAAATATNNACCAGAGAACAACTTTGATCATGTAGACTTCAGCAAAAGGG 601
 OY 89 GlyAsnArgTrpSerIstLeAlaLysHISleuProGlyArgTYrThAspAnGluIleLys 108
 DB 600 GGAACAGGTGTGTCGCAAAATTCGCAAGCATCTACTGTAAGACAGATTAAGATGCAAG 541
 OY 109 AsnTYrTPArGThArgIleGlnIlyshIleLysGlnAlaGluAnSpHegInGln 128
 DB 540 AACTATTGAGAGCCAGATCCAGAGACATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 481
 OY 129 SerSerAsnSerc1u1leAsnAspHISg1nAlaSerTherIstValSerTherMet 148
 DB 480 ATACGATATNACTGTAGATTAATGATACCAAGTACAGCATGTTCTACATG 421
 OY 149 AlaGluProMetGluMetYsrSerProProCysTYrGlnGlyMetLeuGluProPheSer 168
 DB 420 GCTGAACCATGAGACCTATCTCCACCTTTATCAAGGAATGTTAGAGCCATTTCT 361
 OY 169 Thr---GlnPheProThrIleAsnProAspGlnSerSerCysCysThrAsnAspAsn 187
 DB 360 TCATTCACAGTCCCCACCATTAATCTGATCATCATCAGTTGTTGATCAACAAC 301
 OY 188 Asn---IleAsnTYrTrpSerMetGluAspSerTrpSerMetGln-LeuLeuAsnGlyAs 206
 DB 300 AACGATTAATTAATTTGGAGCATGAGGATATCTGTGATATGATTAATGATTAAGAGGGGA 241
 OY 206 P 206
 DB 240 T 240
 RESULT 5
 BM527774 552 bp mRNA linear EST 19-FEB-2002
 LOCUS sal65903.y1 Gm-cl061 glycine max cDNA clone SOYBEAN CLONE ID: 241
 DEFINITION Gm-cl061-4134 5' similar to tr:049020 049020 MYB-LIKE DNA-BINDING
 DOMAIN PROTEIN. ; mRNA sequence.
 ACCESSION BM527774
 VERSION BM527774.1 GI:18733722
 KEYWORDS EST.
 SOURCE soybean.

```

ORGANISM      Glycine max
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                Glycine.
AUTHORS        1 (bases 1 to 552)
                Shoemaker, R., Kelm, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna
                A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
                Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
                Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
                R., Ralter, E., Kohn, S., Shiu, T., Jackson, Y., Cardenas, M., McCann
                R., Waterston, R. and Wilson, R.
                Public Soybean EST Project
                Unpublished (1999)
TITLE          Public Soybean EST Project
JOURNAL        Contact: Shoemaker R./Public Soybean EST Project
                Public Soybean EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available through: Resgen, Invitrogen Corp. 2130
                South Memorial Parkway Huntsville, AL 35801 For further information
                call: (800)-533-4363 or contact: ccuteresgen.com web site:
                www.resgen.com
                Seq primer: -40RP from Gibco
                High quality sequence stop: 426.
FEATURES       Location/Qualifiers
                source          1..552
                /organism="Glycine max"
                /db_xref="taxon:3847"
                /clone="SOYBEAN CLONE ID: Gm-c1061-4134"
                /clone_1lb="Gm-c1061"
                /tissue_type="mature flowers of field grown plants"
                /lab_host="DH10B"
                /note="vector: p Bluescript II SK+, Site_1: EcoRI, Site_2:
                XhoI; The cDNA library was constructed from mRNA isolated
                from mature flowers of field grown plants for the cultivar
                Raiden. Complementary DNA was synthesized from mRNA using
                a primer consisting of a poly(dT) sequence with a XhoI
                restriction site. EcoRI adapters were ligated to the
                blunt-ended cDNA fragments followed by XhoI digestion. The
                cDNA fragments were directionally cloned into the
                pCotRI-XhoI restriction site of the pBluescript vector. The
                ligated cDNA fragments were transformed into DH10B host
                cells (GibcoBRL). This library was constructed in the
                laboratory of Dr. Randy Shoemaker."
BASE COUNT     184 a      118 c      111 g      139 t
ORIGIN
Alignment Scores:
Pred. No.:      1.05e-85      Length:      552
Score:          816.00      Matches:      152
Percent Similarity: 95.62%      Conservative: 1
Best Local Similarity: 95.00%      Mismatches: 5
Query Match:    71.70%      Indels:      2
                Gaps:      2
US-10-021-811-36 (1-206) x BM527774 (1-552)
QY      49 leulysarGAnsglyysSerCysArgleuArgTPluAsnTyrleuArgProAspAl 68
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
        1 CTCAAAGCGTACGGAAGAGTTCGCGGTAAAGTGGCTTAACTACCTCCGTCCTGATGT 60
QY      69 ArgArgGlyAsnIleThrProGluGluGlnIleuIleMetGluLeuHisAlaTyrTrp 88
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
        61 AGAAGAGGGAATATATACCCCGAGGAACAATTTCATCATGACCTTCACCAAGTGG 120
QY      89 GlyAsnArgTyrSerIleAlaLysHisLeuProGlyArgThrAspAsnGluLeuLys 108
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
        121 GGAAGAGGAGTGTCAAAATTCGCAAGCATCTACCTGCTAGAGCAAGATTAATGATCAAG 180
QY      109 AsnTyrTrpArgThrArgIleGlnLysHisIleLysGlnAlaGluAsnPhenGlnGln 128
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```


4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntville, AL 35801 for further information
 call: (800)-533-4363 or contact: ccuteresgen.com web site:
 www.resgen.com

Seq primer: -40RP from gibco
 High quality sequence stop: 422.

FEATURES

source

Location/Qualifiers
 1..568
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-c1061-4003"
 /clone_lib="Gm-c1061"
 /tissue_type="Mature flowers of field grown plants"
 /lab_host="DH10B"

/note="Vector: Bluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from mature flowers of field grown plants for the cultivar Raiden. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 174 a 134 c 124 g 136 t
 ORIGIN

Alignment Scores:

Score: 2.85e-83 Length: 568
 Percent Similarity: 795.50 Matches: 153
 Best Local Similarity: 87.91% Conservative: 7
 Query Match: 69.90% Mismatches: 19
 Gaps: 3
 Indels: 3

US-10-021-811-36 (1-206) x BM527664 (1-568)

OY 1 Metaspyslysginglncyslysthrserclasprrrogluvalarglyglyprotrp 20
 |||||
 DB 29 ATGACCAAGAGCTTGGC--AACACGCTCATATCTCGAAGTGAAGAGGCGCATGG 85
 OY 21 ThmetglugluaspLeuileleumetasnryrillealaasnhsiglyuglyvaltrp 40
 |||||
 DB 86 ACAATGGAAAGACTTAATCTTGATCACCCTATTCACACGGGGAAGGGGTTTGG 145
 OY 41 AsnsertleuialysAlaialglyleuylsrgasnlylysserasyargleuargtrp 60
 |||||
 DB 146 AACCTTTGGCCAAAGCGTGGCTGACCTTAACGTCACCGAAGACTTCCCGCTCCGTCG 205
 OY 61 LeuasnryrleuargproaspvalargarglyasnliethrprogluglucInleuenu 80
 |||||
 DB 206 CTAACCTACCTCCCTCCGATGTAGAAAGGAGATATTACCCCGAGGAACACCTTTTG 265
 OY 81 llemetgluleuhsAlaialystripclYasnargtrpserlylAlaialysHsleuPro 100
 |||||
 DB 266 ATCAGGAACCTTCATCCAAAGTGGGAAACAGCGTCCAAATTTCCCAACCATCTACCC 325
 OY 101 Glyargthraspansgluilelyshsnryrtrpargthrarqllleglnlyshsillelys 120
 |||||
 DB 326 GGAAGAGCTGATTAATGACATTAGAACTACGAGAGAACAGATCCAGAACACCTCAAG 385
 OY 121 Glnala---GluasnPhgInglnglnserSerasnSerGluileasnAspHisGln 139
 |||||
 DB 386 CAAGCTTCAGAGAGCTTCAGACAGAGCTAGTAATTCTGAGATAATTATCATCCCAAA 445
 OY 140 AlaSerThrSerHsvalSerThmetlaclupProkmetglumetYrSerProProCys 159
 |||||

DB 446 GCTTGCAGTACCAAGCTTCACCATGGCGACGCCCATAGAAACCTATTCTCCACCCAGT 505
 OY 160 TyrginglymetleugluProPheserThrglnpneprohrlleasnProaspGlnser 179
 |||||
 DB 506 TATCAAGGAATGTATGATTCATTTTCATTCAGTCCACACA---AATCCATCATCATCT 562
 OY 180 SerCys 181
 |||||
 DB 563 ACTGT 568

RESULT 10

A1930997

LOCUS

523 bp mRNA linear EST 30-NOV-2001
 sb45h07.v1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-c1015-278 5' similar to TR:Q39028 Q39028 ATMVB2. ;, mRNA
 sequence.

ACCESSION

A1930997.1 GI:5666961

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 523)

Shoemaker, R., Kelm, P., Vodka, L., Erpellding, J., Corvelli, V., Khana

A., Bolla, B., Matra, M., Hillier, L., Kucaba, I., Matlin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk

R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 for further information

call: (800)-533-4363 or contact via email: ccuteresgen.com

Seq primer: -40RP from gibco

High quality sequence stop: 422.

FEATURES

source

1..523

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-278"

/clone_lib="Gm-c1015"

/tissue_type="Mature flowers, field grown plants"

/lab_host="X10-Gold"

/note="Vector: pluescript II Xr; Site_1: EcoRI; Site_2:

XhoI; This cDNA library was constructed from mRNA isolated

from mature flowers of field grown plants. The cDNA

library was prepared using the Stratagene pluescript II

XR cDNA library construction kit. Complementary DNA was

synthesized from mRNA using a primer consisting of a poly

(dT) sequence with a XhoI restriction site. EcoRI adapters

were ligated to the blunt-ended cDNA fragments followed by

XhoI digestion. The cDNA fragments were directionally

cloned into the EcoRI-XhoI restriction site of the

pluescript vector. The ligated cDNA fragments were

transformed into X10-Gold host cells. This library was

constructed by Dr. Randy Shoemaker and Dr. John

Expelding."

BASE COUNT 172 a 124 c 111 g 115 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:

5.68e-82

Length:

523

Score: 784.00 Matches: 146
 Percent Similarity: 94.848
 Best Local Similarity: 94.198
 Query Match: 68.898
 DB: 9
 Gaps: 0

US-10-021-811-36 (1-206) x A1930997 (1-523)

QY 1 Metasplyslysglncglnyslthrsrserglnasprglnvalarglyslprotrp 20
 |||||
 Db 58 ATGGAATAAAAACACTGTCACACGCTCTCAAGATCCCGAAGTGAAGAAAGACCTTGG 117
 QY 21 Thmetcgluaspneuileuemetasnyrlllealasnhsiglygluylvaltrp 40
 |||||
 Db 118 ACGATGAAATTAAGACTTGTGATCTGATCACTATATTGCAAAATCATGGGAAAGTGTGG 177
 QY 41 Asnserleualalysalaalaglyleulyasnglylvserscysarglleuartrp 60
 |||||
 Db 178 AATTCCTTGGCCAAACGCTGCTCTCAAAACCTACCGAAAGAGTGGCGGCTAAGCTGG 237
 QY 61 Leuasnrytleuarproasvalargarglyasnlllethrprogluugluinleu 80
 |||||
 Db 238 CTAACACTCTCGCTCTGATGTAGAAAGGCAATATTACACCGAGAAACACTTTTG 297
 QY 81 Ilewetgluleuhsalalysrtpglasnartrpserlyslllealalyslsleupro 100
 |||||
 Db 298 ATCATGGAGCTTCACGCAAAAGTGGGAAACAGTGTGTCAAATATGCCAAGCATCTACT 357
 QY 101 Glyargthrpsangluileulyasnyrtrpartrpartrllleghlyshlslelys 120
 |||||
 Db 358 GGTGAGCAATATGATGATCAAGAACTATGAGAGCAGACGATCCAGAGACATCAAG 417
 QY 121 Glnalagluasnphenglnglnserserasnhsersglulleasnhsiglnala 140
 |||||
 Db 418 CAAGCTGAGAACTGTCAGCAAAATCAGCAATTAATCTGAGATTAATGATCACCAGCT 477
 QY 141 Serthrserhsvalserthrmetalaagluuprometglumetyr 155
 |||||
 Db 478 AGCACTAGCCATGTTCTTACATGCTGAACCATGAGACCTAT 522

RESULT 11 688 bp mRNA linear EST 18-JUL-2001
 LOCUS B1272897
 DEFINITION NF091A12FL1088 Developing flower Medicago truncatula cDNA clone
 ACCESSION NF091A12FL 5', mRNA sequence.
 VERSION B1272897
 KEYWORDS B1272897.1 GI:14882621
 SOURCE EST.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
 Medicago.
 1 (bases 1 to 688)
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Iman, J.T., Weller, J.W. and May, G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula flower library
 Unpublished (2001)
 Contact: May GD
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2310 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7391
 Fax: 580 221 7380
 Email: gmay@noble.org
 Insert Length: 688 Std Error: 0.00
 Plate: 091 row: A column: 12
 Seq primer: TCACACAGGAAACAGCTATGAC.
 Location/Qualifiers
 1..688
 /organism="Medicago truncatula"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

/db_xref="taxon:3880"
 /clone="NF091A12FL"
 /clone_lib="Developing flower"
 /tissue_type="Developing flowers"
 /dev_stage="Developmentally pooled. Contains a mixture of
 very young, developing, fully-opened flowers and flowers
 in early transition into pods."
 /note="Vector: lambda Zap; cDNA was prepared from polyA+
 enriched, pooled samples of equivalent amounts of total
 RNA from very young, developing, fully-opened flowers and
 flowers transitioning into pods. The cDNA was
 directionally ligated into the Uni-Zap XR vector
 (Stratagene) and packaged using the Gigapack III Gold
 packaging extracts. Phagemids containing Uni-Zap XR vector
 in vivo excised from the recombinant Uni-Zap XR vector
 using Exsist helper phage and the E. coli strain
 XL1-Blue MRF' (Stratagene). Excised plasmids were plated
 using SOLR cells."

BASE COUNT 235 a 144 c 125 g 182 t 2 others
 ORIGIN

US-10-021-811-36 (1-206) x B1272897 (1-688)

QY 1 Metasplyslysglncglnyslthrsrserglnasprglnvalarglyslprotrp 20
 |||||
 Db 49 ATGGAATAAAAA---GAGTCACACTCTCTCAAGATCCCGAAGTGAAGAAAGACCATG 105
 QY 21 Thmetcgluaspneuileuemetasnyrlllealasnhsiglygluylvaltrp 40
 |||||
 Db 106 ACAATGGAAGACACTGATCTGATCAACTATATTGCCAATCATGGGAAAGCTGTGG 165
 QY 41 Asnserleualalysalaalaglyleulyasnglylvserscysarglleuartrp 60
 |||||
 Db 166 AATTCCTTGGCCAAACGCTGCTCTCAAAACCTACCGAAAGAGTGGCGGCTGAGGG 225
 QY 61 Leuasnrytleuarproasvalargarglyasnlllethrprogluugluinleu 80
 |||||
 Db 226 CTAACACTCTCTGACCGGATGTAGACGAGGAAATTAATTAACCTGAGAAACACTCTTG 285
 QY 81 Ilewetgluleuhsalalysrtpglasnartrpserlyslllealalyslsleupro 100
 |||||
 Db 286 ATCATTAACCTATGCTGTAAGTGGGAAACAGTGTGCAAAATATGCCAAGCATCTACCG 345
 QY 101 Glyargthrpsangluileulyasnyrtrpartrpartrllleghlyshlslelys 120
 |||||
 Db 346 GGAAGAACCAATATGATGATTAAGAACTACTGAGAGCAAGGATCCAAAGCATATCAAG 405
 QY 121 Glnalagluasnphenglnglnserserasnhsersglulleasnhsiglnala 139
 |||||
 Db 406 CAACCTGAGAACTGCAATCTCAAGAAACAGTTCGATATTCATTAATTAATGATATAT 465
 QY 140 -----Alasethrserhsvalserthrmetalaagluuprometglu 153
 |||||
 Db 466 AATAATATCAACTTATAGGACCAACCAATATTCACAACTTACACCTATGAGC 525
 QY 154 -----Metysrserprocytyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyrtyr 169
 |||||
 Db 526 ACAAT 585
 QY 170 Glnpheprothrllleasnproaspglnserserlyslllealalyslsleupro 189
 |||||
 Db 586 CAGTTCCTCAAACTCT---GATCAATCAGGTGTGCT----- 621
 QY 190 Asnrytrpsermetgluasprsertrpsermetgluileuasnnglyasp 206
 |||||


```

OY 116 Glnlyshis11elysglnalaglualsnphednglnInSerSerAsnAsnSergluile 135
DB 309 CNAACACGACATTAAAGACAGACATGATGACAGAAAGATTACAGACAAATATAT 368
OY 136 AsnAspHisGlnAlaSerThrSerHisValSerThrMet-----AlaGluProMetGlu 153
DB 369 GATGATCAACAGAGAGTACTGATCAATATATCATCTGTTGGTCAACACATCAATACATAGAG 428
OY 154 MetTyrSerProProcTyrGlnGluMetLeuGluProPheSerThrGlnPheProThr 173
DB 429 ACTTACTCTCACTCATCATGATGAAATTTGGAC-----ACTATATTT----- 473
OY 174 IleAsnProAspGlnSerCysCysThrAsnAspAsnAsnIleAsnIleAsnIleAsnIle 193
DB 474 -----CAGGCTCAAAATTTCTCATGAAACAATATAC---AACATGTGGAGC 518
OY 194 MetGluAspSerTyrPheSerMetGlnLeuAsnGlyAsp 206
DB 519 ATGAGAGATATCTGTCATGCAATGCTTAATGGAGAT 557

RESULT 15
LOCUS BE057370 476 bp mRNA linear EST 03-DEC-2001
DEFINITION sn02c04.y1 Gm-cl015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl015-8383 5' similar to TR:049020 049020 M1B-LIKE DNA-BINDING
DOMAIN PROTEIN. ; mRNA sequence.
ACCESSION BE057370
VERSION BE057370.1 GI:8401736
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 476)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpeliding, J., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wylie, T., Underwood, K., Stepien, M., Rheisinger, B., Allen, M., Bowers
J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Rittler, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: cu@resgen.com
Insert Length: 909 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 420.
Location/Qualifiers
1. 476
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl015-8383"
/clone_1lb="Gm-cl015"
/issue_type="Mature flowers, field grown plants"
/lab_host="X110-Gold"
/note="Vector: pluescript II XR; Site_1: EcoRI; Site_2:
XhoI; This cDNA library was constructed from mRNA isolated
from mature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pluescript II
XR cDNA library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by

```

```

XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pluescript vector. The ligated cDNA fragments were
transformed into X110-gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpeliding."
BASE COUNT 145 a 113 c 116 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 2.9e-73 Length: 476
Score: 709.50 Matches: 133
Percent Similarity: 90.85% Conservative: 6
Best Local Similarity: 86.93% Mismatches: 13
Query Match: 62.35% Indels: 1
DB: Gaps: 10
US-10-021-811-36 (1-206) x BE057370 (1-476)
OY 9 ThrSerGlnAspProGluValAlaArgLysGlyProThrMetGluGluAspLeuIleu 28
DB 16 ACGTCTCATGATCCTCGAAGTGAAGAGGGCCATGACAAATGAAAGAACTTAATCTTG 75
OY 29 MetAsnTyrIleAlaAsnHisGlyGluGluValTyrPheSerIleuAlaValAlaGly 48
DB 76 ATCACTATATTTGCCAATCACGGGAGAGGGCTTGGAACCTTTGGCCAAAGCTGCTGGA 135
OY 49 LeuLysArgAsnGlyLysSerCysArgLeuArgTyrPheAsnTyrLeuArgProAspVal 68
DB 136 CTTAAACGTACCGGAAAGAGTTGCCGCTCCGCTTAACCTCCGCTCGATGTTT 195
OY 69 ArgArgGlyAsnIleThrProGluGluGlnLeuIleMetGluIleuAlaLysTyr 88
DB 196 AAGAGAGGAATTTACACCCGAGAGACAGCTTTGATCATGGAACCTTCATGCAAAAGTGG 255
OY 89 GlyAsnArgTyrSerLysIleAlaLysHisLeuProGlyArgThrAspAsnGluIleLys 108
DB 256 GGAACACAGGTGTCCTCAAAATTTCCAAAGCATCTACCCGAGAGAGCTGATTAAGATTAAAG 315
OY 109 AsnTyrTyrArgThrArgIleGlnLysHisIleLysGlnAla---GluAsnPhedngln 127
DB 316 AACTACTGAGAGCAACAAGATCCAGAAGCACATCAAGCAAGCTTCCAGAGCTTCAGCAA 375
OY 128 GlnSerSerAsnAsnSergluIleAsnAspHisGlnAlaSerThrSerHisValSerThr 147
DB 376 CAGAGAGTAGTAATCTGAGATTAATTATCATCCCAAGCTTGCGATGAGCCAGAGTGCACG 435
OY 148 MetAlaGluProMetGluMetTyrSerProProcTyr 160
DB 436 ATGGCCACGCCATACAAACCTATTTCTCCACCCAGCTTAT 474

```

Search completed: February 18, 2003, 11:43:25
 Job time : 1691 secs